AD_____

Award Number: DAMD17-98-1-8305

TITLE: Genomic Imprinting of the M6P/IGF2 Receptor: A Novel Breast Cancer Susceptibility Mechanism

PRINCIPAL INVESTIGATOR: Randy L. Jirtle, Ph.D.

CONTRACTING ORGANIZATION: Duke University Medical Center
Durham, North Carolina 27710

REPORT DATE: July 2001

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2001	Final (1 Jul 98 - 30 Jun 01)

4. TITLE AND SUBTITLE
Genomic Imprinting of the M6P/IGF2 Receptor: A Novel Breast Cancer Susceptibility Mechanism

5. FUNDING NUMBERS
DAMD17-98-1-8305

6. AUTHOR(S)
Randy L. Jirtle, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Duke University Medical Center
Durham, North Carolina 27710

E-Mail: jirtle@radonc.duke.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

Genomic imprinting is an epigenetic phenomenon in mammals that results in the differential expression of the paternally and maternally inherited alleles of a gene. Imprinted genes normally function to control embryonic growth and development. They also are involved in cancer because their functional haploid state makes them vulnerable to being either inactivated or overexpressed. The M6P/IGF2R has been shown to suppress cancer cell growth and is mutated in a number of human cancers, including those that develop in the lung, liver, colon and breast. These findings are consistent with the *M6P/IGF2R* functioning normally as a tumor suppressor. We have shown that *M6P/IGF2R* imprinting and receptor IGF2 binding evolved in an ancestor common to marsupials and eutherian mammals. Although *M6P/IGF2R* is imprinted in lower eutherian mammals, we have demonstrated that it is not imprinted in humans, clarifying previous misconceptions in the literature regarding its imprint status. Our results with breast cancer and Wilms tumor suggest that a mutational event within intron 10 of the *M6P/IGF2R*, and not disregulation of imprinting, may have resulted in the aberrant monoallelic expression observed in some humans. Therefore, although the *M6P/IGF2R* normally functions as a tumor suppressor in lung, liver, colon and breast cancer, genomic imprinting at this locus is not involved in the etiology of tumorigenesis.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast cancer, genomic imprinting, M6P/IGF2R, tumor suppressor, imprinting evolution	27
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

TABLE OF CONTENTS

INTRODUCTION

The *mannose 6-phosphate/insulin-like growth factor II receptor (M6P/IGF2R)* gene encodes for a receptor that plays a critical role in regulating the bioavailability of extracellular proteolytic enzymes and growth factors known to be involved in carcinogenesis (*1,2*). Our recent findings indicate that the *M6P/IGF2R* also functions as a tumor suppressor gene in liver, breast, and lung cancer (*2,3,4,5*). We have determined that the frequency of monoallelic *M6P/IGF2R* expression in breast cancer patients is higher than that of age-matched controls. However, we have also demonstrated that the *M6P/IGF2R* is not imprinted in humans (*6*). Therefore, the observed monoallelic *M6P/IGF2R* expression in breast cancer cannot be attributed to aberrant imprint regulation.

BODY

Genomic imprinting is a non-Mendelian, parent-of-origin inherited, epigenetic form of gene regulation that results in monoallelic expression. The *M6P/IGF2R* is imprinted in both rats and mice, but the imprint status in humans at this locus has been reported to be polymorphic (*7,8*; for review see *2,4,9,10,11*). Because of this species difference in *M6P/IGF2R* imprinting, rodents would be predicted to be more sensitive than humans to cancer because only one allele would need to be mutated to inactivate its tumor suppressor function (*2,10*). Therefore, it is important to better understand the phenomenon of genomic imprinting, and its modification by both genotoxic and non genotoxic agents since rodents are used as surrogates for human cancer risk assessment (*11,12*).

The literature reports of *M6P/IGF2R* imprinting suggest that this gene may be polymorphically imprinted in humans, with some individuals expressing only the maternal allele and most other individuals expressing both parental alleles. Because of this uncertainty, and also due to our previous finding in Wilms tumor patients with monoallelic upstream *M6P/IGF2R* expression and biallelic downstream *M6P/IGF2R* expression, we wished to establish whether the *M6P/IGF2R* is subject to genomic imprinting. For some genes, imprinting can be age-dependent, with imprinted expression occurring only during early development. Therefore, to determine if the human *M6P/IGF2R* is imprinted, we utilized organ tissues derived from 75 human conceptuses (40 first trimester and 35 second trimester). Using six polymorphisms within the *M6P/IGF2R* that were discovered in our laboratory (*13*), all informative conceptuses (46/75; 61%) were shown to express *M6P/IGF2R* biallelically (*6*). This analysis showed that the *M6P/IGF2R* is not subject to genomic imprinting during fetal development. Recent additional experimental evidence from our laboratory, based on a detailed analysis of the evolution of imprinting of the *M6P/IGF2R* indeed supports that the *M6P/IGF2R* is not imprinted in humans (6,*14*).

Genomic imprinting is postulated to have evolved because of a parent-offspring conflict to control fetal growth. This parental "tug-of-war" model predicts that only eutherian mammals would have imprinted genes because of the intrauterine development

of their offspring. We tested this postulate by comparing *M6P/IGF2R* imprinting in monotremes (i.e. echidna and platypus), marsupials (i.e. opossum) and eutherian mammals (i.e. mouse, rat, pig, cow, bat, flying lemur, tree shrew, ringtail lemur and humans). Our findings demonstrate that *M6P/IGF2R* is not imprinted in the egg-laying platypus and echidna, whereas it is imprinted in the opossum (*14*). Thus, imprinting evolved in viviparous mammals over 100 million years ago; however, since the opossum lacks a fetal stage of development, invasive placentation and intrauterine fetal growth are not required for genomic imprinting to evolve. The *M6P/IGF2R* in both the monotremes and didelphid marsupials also lacks the differentially-methylated CpG island in intron 2 previously postulated to be mechanistically involved in imprint control in mice. This demonstrates the existence of alternative mechanisms of *M6P/IGF2R* imprint establishment and maintenance. Our results also indicate that monotremes and marsupials are not as closely related as predicted by the Marsupionta model; instead, they support the morphology-based Theria hypothesis of mammalian evolution (*15*).

We have also shown that although the *M6P/IGF2R* is imprinted in mice, rats, pigs, cows, and bats, imprinting at this locus was lost approximately 70 million years ago with the evolution of the higher mammalian orders: Dermoptera (e.g. flying lemurs), Scandentia (e.g. tree shrews), and Primates (e.g. ringtail lemurs and humans) (*6*). This finding provides compelling evidence that *M6P/IGF2R* imprinting is not a polymorphic trait in humans since convergent evolution of *M6P/IGF2R* imprinting would have had to have occurred in humans. This highly unlikely possibility is also supported by our inability to demonstrate *M6P/IGF2R* imprinting in either fetal or adult human tissues. The lack of *M6P/IGF2R* imprinting in humans has important ramifications in toxicology because it strongly indicates that although the *M6P/IGF2R* functions as a tumor suppressor in humans and rodents, rodents are at heightened susceptibility to tumor formation because of the imprinted status and consequent functionally haploid state of the *M6p/Igf2r*.

KEY RESEARCH ACCOMPLISHMENTS

- *M6P/IGF2R* is frequently mutated in human breast, liver and lung cancer suggesting it functions as a tumor suppressor gene.
- Monoallelic *M6P/IGF2R* 3' end gene expression was found in 2/32 (6.3%) of breast cancer patients, suggestive of a posttranscriptional mechanism which results in the production of a single or truncated mRNA species.
- In Wilms tumor patients, a truncated *M6P/IGF2R* transcript is produced from one allele with the site of truncation within intron 10 of the *M6P/IGF2R* gene.
- *M6P/IGF2R* is not imprinted in humans.
- *M6P/IGF2R* imprinting (i.e. maternal expression) and receptor IGF2 binding evolved in an ancestor common to marsupials and eutherian mammals.

- The evolutionary loss of imprinting of the *M6P/IGF2R* in higher mammals (Dermoptera, Scandentia, and Primates) approximately 70 million years ago supports that the *M6P/IGF2R* is not imprinted in humans.

- The finding that the *M6P/IGF2R* is not imprinted in humans indicates that imprinting is not a contributing factor in the etiology of breast cancer and Wilms tumor.

REPORTABLE OUTCOMES

- Killian, J.K., and Jirtle, R.L. Genomic structure of the human *M6P/IGF2* receptor. ***Mamm. Genome*** 10: 74-77, 1999.
- Falls, J.G., Wylie, A.A., Pulford, D.J., and Jirtle, R.L. Genomic imprinting: implications in human disease. ***Am. J. Pathol.*** 154: 635-647, 1999.
- Jirtle, R.L. Genomic imprinting and cancer. ***Exp. Cell Res.*** 248: 18-24, 1999.
- Devi, G.R., De Souza, A.T., Byrd, JC, Jirtle. R.L., and MacDonald, R.G. Altered ligand binding by insulin-like growth factor II/mannose 6-phosphate receptors bearing missense mutations in human cancers. ***Cancer Res.*** 59: 4314-4319, 1999.
- Byrd, J.C., Devi, G.R., De Souza, A.T., Jirtle, R.L., and MacDonald, R.G. Disruption of ligand binding to the insulin-like growth factor II/mannose 6-phosphate receptor by cancer-associated missense mutations. ***J. Biol. Chem.*** 274: 24408-24416, 1999.
- Jirtle, RL. Mannose 6-Phosphate Receptors. In: Encyclopedia of Molecular Biology, (Creidton, T.E., ed.), pp. 1441-1447, Wiley-Liss, Inc., New York, New York, 1999.
- Jirtle, R.L., Sander, M., and Barrett, J.C. Genomic imprinting and environmental disease susceptibility. ***Environ. Health Perspect.*** 108: 271-278, 2000.
- Murphy, S.K., and Jirtle, R.L. Imprinted genes as potential genetic and epigenetic toxicologic targets. ***Environ. Health Perspect.*** 108 (Suppl 1): 5-11, 2000.
- Killian, J.K., Byrd, J.C, Jirtle, J.V., Munday, B.L., Stoskopf, M.K., and Jirtle, R.L. *M6P/IGF2R* imprinting evolution in mammals. ***Mol. Cell*** 5: 707-716, 2000.
- Kong, F., Anscher, M.S., Washington, M.K., Killian, J.K., Jirtle, R.J. *M6P/IGF2R* is mutated in squamous cell carcinoma of the lung. ***Oncogene*** 19:1572-1578, 2000.
- Wylie, A. A., Murphy, S. K., Orton, T. C., Jirtle, R. L. Regulatory Motifs of the Novel Imprinted Domain, *DLK1/GTL2*, Mimic Those of *IGF2/H19*. ***Genome Res.*** 10:1711-1718, 2000.
- Murphy, S.K., Wylie, A.A., Jirtle, R.J. Imprinting of *PEG3*, the human homolog of a gene involved in nurturing behavior. ***Genomics*** 71(1):110-117, 2001.
- Evans, H.K., Wylie, A. A., Murphy, S. K., Jirtle, R.L. *NNAT* resides in a micro-imprinted domain on human chromosome 20q11.2. ***Genomics***. In press.
- Killian, J.K., Oka, Y., Jang, H., Fu, X., Waterland, R.A., Sohda, T., Sakaguchi, S., Jirtle, R.L. Mannose 6-phosphate/insulin-like growth factor 2 receptor (M6P/IGF2R) variants in American and Japanese populations. ***Hum. Mutat.*** 18, 25-31, 2001.
- Nolan, C.M., Killian, J.K., Petitte, J.N., Jirtle, R.L. Imprint status of *M6P/IGF2R* and *IGF2* in chickens. ***Dev. Genes Evol.*** 211:179-183, 2001.
- Killian, J.K., Buckley, T.R., Stewart, N., Munday, B.L., Jirtle, R.L. Marsupials and Eutherians reunited: genetic evidence for the Theria hypothesis of mammalian evolution. ***Mamm. Genome*** 12:513-517, 2001.

- Murphy, S. K., Jirtle, R.L. Non-genotoxic Causes of Cancer. In: *The Cancer Handbook* (Alison, M.R., and N. G. Gooderham, eds). Nature Publishing Group Reference, London. In press.
- Killian, J.K., Li, T., Nolan, C.M., Vu, T., Hoffman, A.R., Jirtle, R.L. Divergent evolution in genomic imprinting from the Jurassic to the Quaternary. ***Hum. Mol. Genet.*** In press.
- Killian, J.K., Nolan, C.M., Stewart, N., Munday, B.L., Andersen, N.A., Stewart Nicol, S., Jirtle, R.L. Monotreme IGF2 expression and ancestral origin of genomic imprinting. ***J. Exp. Zoology.*** In press.

CONCLUSIONS

In conclusion, there is now compelling mutational and functional evidence that the *M6P/IGF2R* is a tumor suppressor that is frequently inactivated during the early stages of human cancer formation. *M6P/IGF2R* loss of function not only provides cancer cells with an early growth advantage, but also confers enhanced resistance to radiotherapy treatment. However, the lack of *M6P/IGF2R* imprinting indicates that this gene does not confer susceptibility to human cancer because of genomic imprinting. This important result stresses the need for careful extrapolation of rodent carcinogenic risk assessment data to humans, since rodents are indeed imprinted at this locus and are therefore at heightened susceptibility to tumor formation because of the haploid expression of the M6p/Igf2r. Future studies investigating the tumor suppressive role of the *M6P/IGF2R* in human breast cancer can now focus on disruptions both to the biological functions this important protein serves in tumor cells and the regulatory mechanisms controlling its expression.

REFERENCES

1. S. Kornfeld, *Annu. Rev. Biochem.* **61**, 307 (1992).
2. R.L. Jirtle, *Exp. Cell Res.* **248,** 18 (1999).
3. A.T. DeSouza, T. Yamada, J.J. Mills, R.L. Jirtle, *FASEB J.* **11**, 60 (1997).
4. J.G. Falls, D.J. Pulford, A.A. Wylie, R.L. Jirtle, *Am. J. Pathol.* **154**, 635 (1999).
5. Kong, F., Anscher, M.S., Washington, M.K., Killian, J.K., Jirtle, R.J. *Oncogene* **19**, 1572 (2000).
6. Killian, J.K., Li, T., Nolan, C.M., Vu, T., Hoffman, A.R., Jirtle, R.L. *Hum. Mol. Genet.* In press.
7. Xu, Y., Goodyer, C.G., Deal, C., Polychronakos, C. *Biochem. Biophys. Res. Commun.* **197**, 747 (1993).
8. Xu, Y., Grundy, P., Polychronakos, C. *Oncogene* **14**, 1041 (1997).
9. Jirtle, RL. *In* Encyclopedia of Molecular Biology, (Creidton, T.E., ed.), 1441-1447, Wiley-Liss, Inc., New York, New York, (1999).
10. Pulford, D.P., Falls, J.G., Killian, J.K., Jirtle, R.J. *Mutat. Res.* **436**, 59 (1999).
11. Murphy, S.K., Jirtle, R.L. *Environ. Health Perspect.*. **108 (Supp 1)**, 5 (2000).
12. Jirtle, R.J. *Environ. Health Perspect.* **108**, 271 (2000)

13. Killian, J.K., Oka, Y., Jang, H., Fu, X., Waterland, R.A., Sohda, T., Sakaguchi, S., Jirtle, R.L. *Human Mutation* **18**, 25-31 (2001).
14. Killian, J.K., Byrd, J.C, Jirtle, J.V., Munday, B.L., Stoskopf, M.K., Jirtle, R.L. *Mol. Cell* **5**, 707 (2000).
15. Killian, J.K., Buckley, T.R., Stewart, N., Munday, B., Jirtle, R.L. *Mammalian Genome* **12**, 513-517 (2001).

BIBLIOGRAPHY

1. Byrd, J.C., Devi, G.R., De Souza, A.T., Jirtle, R.L., and MacDonald, R.G. Disruption of ligand binding to the insulin-like growth factor II/mannose 6-phosphate receptor by cancer-associated missense mutations. ***J. Biol. Chem.*** 274: 24408-24416, 1999.
2. Devi, G.R., De Souza, A.T., Byrd, JC, Jirtle. R.L., and MacDonald, R.G. Altered ligand binding by insulin-like growth factor II/mannose 6-phosphate receptors bearing missense mutations in human cancers. ***Cancer Res.*** 59: 4314-4319, 1999.
3. Evans, H.K., Wylie, A. A., Murphy, S. K., Jirtle, R.L. *NNAT* resides in a micro-imprinted domain on human chromosome 20q11.2. ***Genomics***. In press.
4. Falls, J.G., Wylie, A.A., Pulford, D.J., and Jirtle, R.L. Genomic imprinting: implications in human disease. ***Am. J. Pathol.*** 154: 635-647, 1999.
5. Jirtle, RL. Mannose 6-Phosphate Receptors. In: Encyclopedia of Molecular Biology, (Creidton, T.E., ed.), pp. 1441-1447, Wiley-Liss, Inc., New York, New York, 1999.
6. Jirtle, R.L. Genomic imprinting and cancer. ***Exp. Cell Res.*** 248: 18-24, 1999.
7. Jirtle, R.L., Sander, M., and Barrett, J.C. Genomic imprinting and environmental disease susceptibility. ***Environ. Health Perspect.*** 108: 271-278, 2000.
8. Killian, J.K., Buckley, T.R., Stewart, N., Munday, B.L., Jirtle, R.L. Marsupials and Eutherians reunited: genetic evidence for the Theria hypothesis of mammalian evolution. ***Mamm. Genome*** 12:513-517, 2001.
9. Killian, J.K., Byrd, J.C, Jirtle, J.V., Munday, B.L., Stoskopf, M.K., and Jirtle, R.L. *M6P/IGF2R* imprinting evolution in mammals. ***Mol. Cell*** 5: 707-716, 2000.
10. Killian, J.K., and Jirtle, R.L. Genomic structure of the human *M6P/IGF2* receptor. ***Mamm. Genome*** 10: 74-77, 1999.
11. Killian, J.K., Li, T., Nolan, C.M., Vu, T., Hoffman, A.R., Jirtle, R.L. Divergent evolution in genomic imprinting from the Jurassic to the Quaternary. ***Hum. Mol. Genet.*** In press.
12. Killian, J.K., Nolan, C.M., Stewart, N., Munday, B.L., Andersen, N.A., Stewart Nicol, S., Jirtle, R.L. Monotreme IGF2 expression and ancestral origin of genomic imprinting. ***J. Exp. Zoology.*** In press.
13. Killian, J.K., Oka, Y., Jang, H., Fu, X., Waterland, R.A., Sohda, T., Sakaguchi, S., Jirtle, R.L. Mannose 6-phosphate/insulin-like growth factor 2 receptor (M6P/IGF2R) variants in American and Japanese populations. ***Hum. Mutat.*** 18, 25-31, 2001.
14. Kong, F., Anscher, M.S., Washington, M.K., Killian, J.K., Jirtle, R.J. *M6P/IGF2R* is mutated in squamous cell carcinoma of the lung. ***Oncogene*** 19:1572-1578, 2000.
15. Murphy, S.K., and Jirtle, R.L. Imprinted genes as potential genetic and epigenetic toxicologic targets. ***Environ. Health Perspect.*** 108 (Suppl 1): 5-11, 2000.
16. Murphy, S. K., Jirtle, R.L. Non-genotoxic Causes of Cancer. In: *The Cancer Handbook* (Alison, M.R., and N. G. Gooderham, eds). Nature Publishing Group Reference, London. In press.
17. Murphy, S.K., Wylie, A.A., Jirtle, R.J. Imprinting of *PEG3*, the human homolog of a gene involved in nurturing behavior. ***Genomics*** 71(1):110-117, 2001.
18. Nolan, C.M., Killian, J.K., Petitte, J.N., Jirtle, R.L. Imprint status of *M6P/IGF2R* and *IGF2* in chickens. ***Dev. Genes Evol.*** 211:179-183, 2001.
19. Wylie, A. A., Murphy, S. K., Orton, T. C., Jirtle, R. L. Regulatory Motifs of the Novel Imprinted Domain, *DLK1/GTL2*, Mimic Those of *IGF2/H19*. ***Genome Res.*** 10:1711-1718, 2000.

PERSONNEL

1. Susan K. Murphy, Ph.D.
   November 1998 to June 2001

2. Randy L. Jirtle, Ph.D.
   November 1998 to June 2001

Dev Genes Evol (2001) 211:179–183
DOI 10.1007/s004270000132

SHORT COMMUNICATION

Catherine M. Nolan · J. Keith Killian
James N. Petitte · Randy L. Jirtle

Imprint status of *M6P/IGF2R* and *IGF2* in chickens

Received: 21 September 2000 / Accepted: 10 December 2000 / Published online: 14 March 2001

Abstract Genomic imprinting is a method of gene regulation whereby a gene is expressed in a parent-of-origin-dependent fashion; however, it is hypothesized that imprinting should not occur in oviparous taxa such as birds. Therefore, we examined the allelic expression of two genes in the chicken that are reciprocally imprinted in most mammals, *mannose 6-phosphate/insulin-like growth factor 2 receptor (M6P/IGF2R)* and *insulin-like growth factor 2 (IGF2)*. Single nucleotide polymorphisms were identified in these genes, and cDNA was prepared from several tissues of embryos heterozygous for these polymorphisms. Both alleles of *M6P/IGF2R* and *IGF2* were expressed in all tissues examined by RT-PCR. Since the expression of these genes was independent of the parent from which they were inherited, we conclude that neither *M6P/IGF2R* nor *IGF2* are imprinted in the chicken.

Keywords Genomic imprinting · Biallelic expression · Chicken · *IGF2* · *M6P/IGF2R*

Introduction

Genomic imprinting is a phenomenon whereby gene expression is parent-of-origin dependent. It appears to have evolved over 100 million years ago (Killian et al. 2000) and the actively debated genetic conflict hypothesis purports that it resulted from a parental genetic "tug-of-war" to control maternal-dependent growth of mammalian offspring (Moore and Haig 1991). According to this hypothesis, growth of the offspring should be promoted by paternally expressed genes and limited by those that are maternally expressed. Furthermore, the duration and extent of prenatal growth and postnatal care should be major driving forces in the evolution of genomic imprinting.

The reciprocal imprinting of *mannose 6-phosphate/insulin-like growth factor 2 receptor* (*M6P/IGF2R*) and *insulin-like growth factor 2* (*IGF2*) in mice and the phenotypes of mice deficient in their expression provide support for the genetic conflict hypothesis. Paternally expressed *IGF2* encodes for a critical fetal mitogen, and mice deficient in this growth factor have a dwarf phenotype (DeChiara et al. 1990) while mice over-expressing *Igf2* are larger than normal (Leighton et al. 1995). The M6P/IGF2R in viviparous mammals binds IGF2 and phosphomannosyl glycoproteins through independent binding sites and limits the biological activity of IGF2 by targeting the growth factor for degradation in lysosomes (Nolan et al. 1990; Kornfeld 1992; Dahms et al. 1993; Killian et al. 2000). The *M6PR/IGF2R* in mice is expressed only from the maternal allele (Barlow et al. 1991), and animals lacking M6PR/IGF2R accumulate excess IGF2, have somatic overgrowth and die perinatally (Lau et al. 1994; Wang et al. 1994; Ludwig et al. 1996; Nolan and Lawlor 1999).

The genetic conflict hypothesis also predicts that genes that regulate embryonic development would not be imprinted in oviparous taxa because in these animals genes expressed during embryogenesis cannot influence the amount of resources they receive from the mother (Haig and Graham 1991). However, recent work in the chicken examining the imprinted status of *IGF2* has yielded conflicting results. Koski et al. (2000) reported that *IGF2* was expressed monoallelically in some embryos from either paternal or maternal alleles. At the same time O'Neill et al. (2000) reported that *IGF2* was expressed in chick embryos in a strictly biallelic manner. Therefore, we re-examined the allelic expression of chicken *IGF2* and examined expression of chicken *M6PR/IGF2R* for the first

Edited by D. Tautz

C.M. Nolan · J.K. Killian · R.L. Jirtle (✉)
Departments of Radiation Oncology and Pathology,
Duke University Medical Center, Durham NC 27710, USA
e-mail: jirtle@radonc.duke.edu
Tel.: +1-919-6842770, Fax: +1-919-6845584
URL: http://www.geneimprint.com

C.M. Nolan
Department of Zoology, University College Dublin, Dublin, Ireland

James N. Petitte
Department of Poultry Science, North Carolina State University, Raleigh, NC 27695–7608, USA

time, and discovered that expression was biallelic for both genes. Our findings are consistent with the predictions of the genetic conflict hypothesis.

Materials and methods

Fertile chicken eggs (White Leghorn or Barred Plymouth Rock strains) were generated at North Carolina State University Department of Poultry Science by artificial insemination of females and were incubated at 37.5°C and 60% relative humidity. Three-day-old embryos were dissected free of membranes, frozen immediately on dry ice and stored at –80°C. Individual tissues (i.e. liver, lungs, heart, skeletal muscle and brain) were removed from 17-day-old embryos and treated similarly. Blood was withdrawn from adult chickens into heparinized syringes and frozen at –80°C.

DNA and total RNA were isolated from whole adult blood or from embryos or tissues using DNA-STAT 120 or RNA-STAT-60 according to the manufacturer's instructions (Tel-test, Friendswood, Tex.). RNA (1–5 μg) was reverse transcribed using Superscript II and oligodT primers, according to the manufacturer's instructions (Life Technologies, Grand Island, N.Y.), and 1/40 of the cDNA obtained was used as a template for PCR amplification. PCR was routinely performed in a 30-μl reaction volume using 1.5 U platinum Taq DNA polymerase (Life Technologies, Baltimore, Md.), 15 pmol primers, 1.5 mM $MgCl_2$, and 100 μM dNTPs (94°C×15 s, 65°C×5 s, and 72°C×60 s for 35 cycles). Amplification products were purified from agarose gels using GenElute spin columns (Sigma, St Louis, Mo.). They were sequenced manually using radiolabeled terminator cycle sequencing (USB Corporation, Cleveland, Ohio) or with an automated ABI 377 sequencer (PE Biosystems, Foster City, Calif.) using the manufacturer's BigDye terminator cycle sequencing kit.

Polymorphic sites were identified in the chicken *M6PR/IGF2R* using genomic DNA amplified with primers designed to putative exon 48, M6PR48F2 (5′-gggaaggagagaa-agatgtcatggttg) and M6PR48R1 (5′-ggaatacccacgtcatatcttttattg-aaattgc); the amplified fragment was sequenced using the ABI 377 sequencer (PE Biosystems, Foster City, Calif.). For analysis of allelic expression of the *M6PR/IGF2R*, cDNA was amplified using M6PR48R1 and a primer located in exon 47 of the gene, M6PR47F1 (5′-ctgcagaagaacatcgggtgttc); the amplified fragment was sequenced using M6PR48R1. Chicken *M6PR/IGF2R* intron 2 was amplified using primers Int2F (5′-tgtggaagatcaagtgcagtct) and Int2R (5′-ggtgctcactgccttgctgg), designed with reference to sequences in exon 2 and exon 3, respectively (Zhou et al. 1995). The Expand Long Template PCR system (Roche Boehringer Mannheim, Indianapolis, Ind.) was used with buffer 3 and amplification conditions of 94°C×10 s, 61°C×20 s, and 68°C×4 min for 30 cycles. The amplified fragment was sequenced using intron-specific primers (sequences available from the authors by request) by a combination of manual and automatic sequencing, and was analyzed by Grail informatics (http://compbio.ornl.gov/grail-bin/EmptyGrailForm). The complete sequence of intron 2 of the chicken *M6P/IGF2R* is deposited in GenBank (Accession No. AF305581).

To identify single nucleotide polymorphisms in chicken *IGF2*, a portion of the gene was amplified using primers IGF2F2 (5′-cctcccagccctcaacaag) and IGF2R2 (5′-tcccc-aggagatcacaaatcgag) and sequenced using the ABI 377 sequencer. These primers span an intron and were also used to amplify *IGF2* cDNA. Allelic expression was analyzed by manual sequencing with the primer IGF2R5 (5′-ccgctgcgagctcttcttc).

Results and discussion

Imprint status of *IGF2* in chickens

IGF2 has been highly conserved in vertebrate evolution and is imprinted in all investigated eutherian and marsupial mammals with only the paternal allele being expressed (DeChiara et al. 1990; Rainier et al. 1993; Overall et al. 1997; Feil et al. 1998; O'Neill et al. 2000). Tissue-specific biallelic expression of *IGF2* does occur, however, in mouse and human brain (Hu et al. 1995; Pham et al. 1998) and adult human liver (Vu and Hoffman 1994). Allelic expression of *IGF2* has also been analyzed in the chicken, with independent studies generating conflicting results (Koski et al. 2000; O'Neill et al. 2000). Further analysis of *IGF2* expression in this species is therefore required.

We identified a T to C transition in exon 3 of the *IGF2* coding sequence of White Leghorn chickens (position 936 of Accession no. AF218827; position 943 of Accession no. S82962). Amplified genomic DNA prepared from skeletal muscle of 17-day-old chicken embryos (n=18) identified six individuals heterozygous for this single nucleotide polymorphism (SNP). RNA and cDNA were then prepared from liver, lungs, heart, skeletal muscle and brain of these informative embryos. Both *IGF2* alleles were found to be expressed in all tissues examined (Fig. 1). We also found biallelic *IGF2* expression in heterozygous 3-day-old embryos (n=17, data not shown). These results are consistent with the conclusion of O'Neill et al. (2000) that *IGF2* is not imprinted in the chicken and agree with the prediction of the genetic

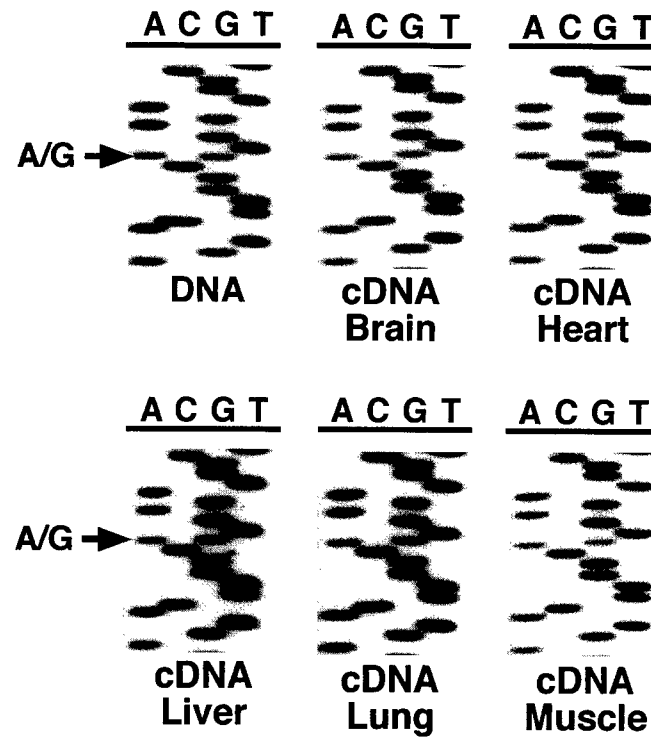

Fig. 1 Allelic expression of the chicken *insulin-like growth factor 2* (*IGF2*). Sequence analysis of genomic DNA amplified from skeletal muscle of a 17-day-old chicken embryo identifies an individual heterozygous for a T to C polymorphism (*arrow*, sequenced in reverse direction). Similar analysis of cDNA prepared from embryonic tissues of the same embryo demonstrates expression of both alleles. The analysis shown is representative of all informative embryos analyzed

conflict hypothesis that genes that regulate embryonic development would not be imprinted in members of oviparous taxa (Moore and Haig 1991).

Koski et al. (2000) also examined *IGF2* expression in chicken embryos and reported that 33% of chickens have monoallelic expression from either the paternal or maternal allele. They concluded that imprinting at the *IGF2* locus is a polymorphic trait. Thus, there is a significant discrepancy between these results and our observations and those of O'Neill et al. (2000). This discrepancy does not result from *IGF2* imprinting being dependent upon stage of development since we found biallelic *IGF2* expression in 3-day-old embryos, the same age studied by Koski et al. (2000). Each of the heterozygote embryos we examined was based on separate individual matings of an out-bred line of chickens and we did not find any evidence for polymorphic *IGF2* imprinting. There are several possible reasons for the observed difference in imprint status of chicken *IGF2*. Apparent monoallelic expression of a non-imprinted *IGF2* gene could result from the preferential PCR amplification of a single allele when template DNA/cDNA concentration is low. It could also result from genetic anomalies such as somatic mosaicism in the parents or loss of heterozygosity at the *IGF2* locus in embryos. Such anomalies do occur in the chicken and karyotypic mosaicism (Abdel-Hameed and Shoffner 1971) and triploidy (Ohno et al. 1963) can be compatible with viable postnatal life in this species. In addition parthenogenetic chickens can develop to maturity (Sarvela 1973). Rather than inferring the genotype of embryos from those of their parents, embryos should be accurately genotyped to guarantee that they are heterozygous at the investigated polymorphic locus. It may be significant that the embryos analyzed by Koski et al. (2000) were not reported as having been genotyped and that monoallelic expression was limited to the offspring of one specific male. Therefore, the random monoallelic expression described by Koski et al (2000) clearly does not constitute the parent-of-origin-dependent expression that is the hallmark of imprinted genes. Taken together with the results of the current study, the available evidence is most consistent with *IGF2* not being imprinted in chickens.

Imprint status of *M6P/IGF2R* in chickens

In viviparous mammals the *M6P/IGF2R* encodes for a receptor that binds both phosphomannosyl glycoproteins and IGF2 through independent binding sites (Kornfeld 1992; Dahms et al. 1993; Yandell et al. 1999). The platypus, chicken and *Xenopus* receptors lack the ability to bind IGF2, however (Clairmont and Czech 1989; Zhou et al. 1995; Killian et al. 2000). It has been proposed that IGF2 binding by M6P/IGF2R evolved in response to high levels of paternally produced IGF2 (Haig and Graham 1991), and that natural selection would have favored inactivating the paternal allele once the receptor acquired its ability to bind IGF2. According to this hypothesis, *M6P/IGF2R* is predicted to be biallelically expressed in the chicken as it is in monotremes (Killian et al. 2000).

Allelic expression of the chicken *M6P/IGF2R* was determined with a G to A transition SNP at nucleotide position 8,637 in the 3′UTR (Accession No. U35037). Embryos heterozygous for the polymorphism were identified by sequencing of genomic DNA isolated from the skeletal muscle of 24 embryos. RNA was then extracted from several tissues of these informative embryos (5 White Leghorn, and 2 Barred Plymouth Rock). RT-PCR was used to amplify the region containing the polymorphism, and the amplified fragment was sequenced to identify expressed alleles. In all heterozygous embryos examined and in all tissues examined (i.e., liver, lungs, heart, skeletal muscle and brain) both alleles of *M6P/IGF2R* were expressed (Fig. 2). We also showed biallelic expression of *M6P/IGF2R* in blood of heterozygous adult chickens (n=21, data not shown). Therefore, we conclude that the chicken *M6P/IGF2R* is not imprinted. Thus, the M6P/IGF2R's acquisition of IGF2 binding, and its resultant ability to limit embryonic growth by degrading IGF2, appears to be closely related to the evolution of imprinting at the *M6P/IGF2R* locus.

The molecular mechanism by which the imprinted status of the *M6P/IGF2R* is achieved and maintained

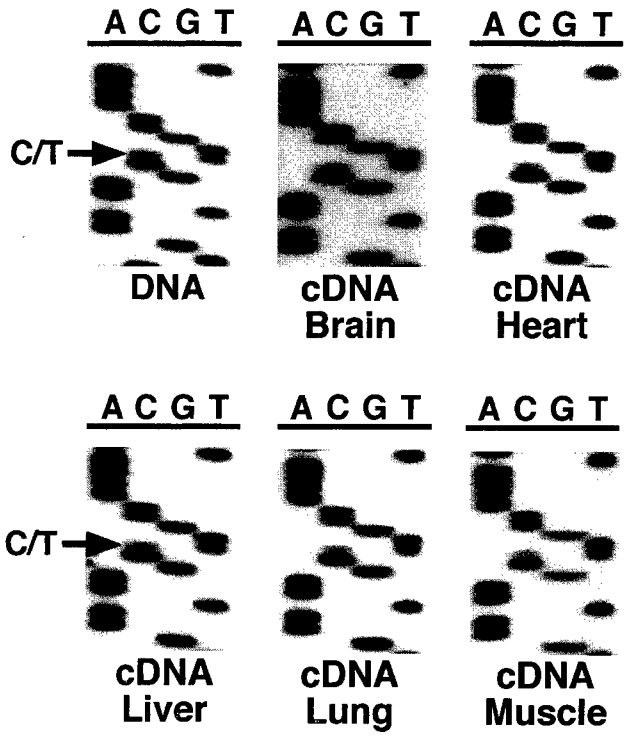

Fig 2 Allelic expression of the chicken *mannose 6-phosphate/insulin-like growth factor 2 receptor* (*M6P/IGF2R*). Sequence analysis of genomic DNA amplified from skeletal muscle of a 17-day-old chicken embryo identifies an individual heterozygous for a G to A polymorphism (*arrow*, sequenced in reverse direction). Similar analysis of cDNA prepared from embryonic tissues of the same embryo demonstrates expression of both alleles. The analysis shown is representative of all informative embryos analyzed

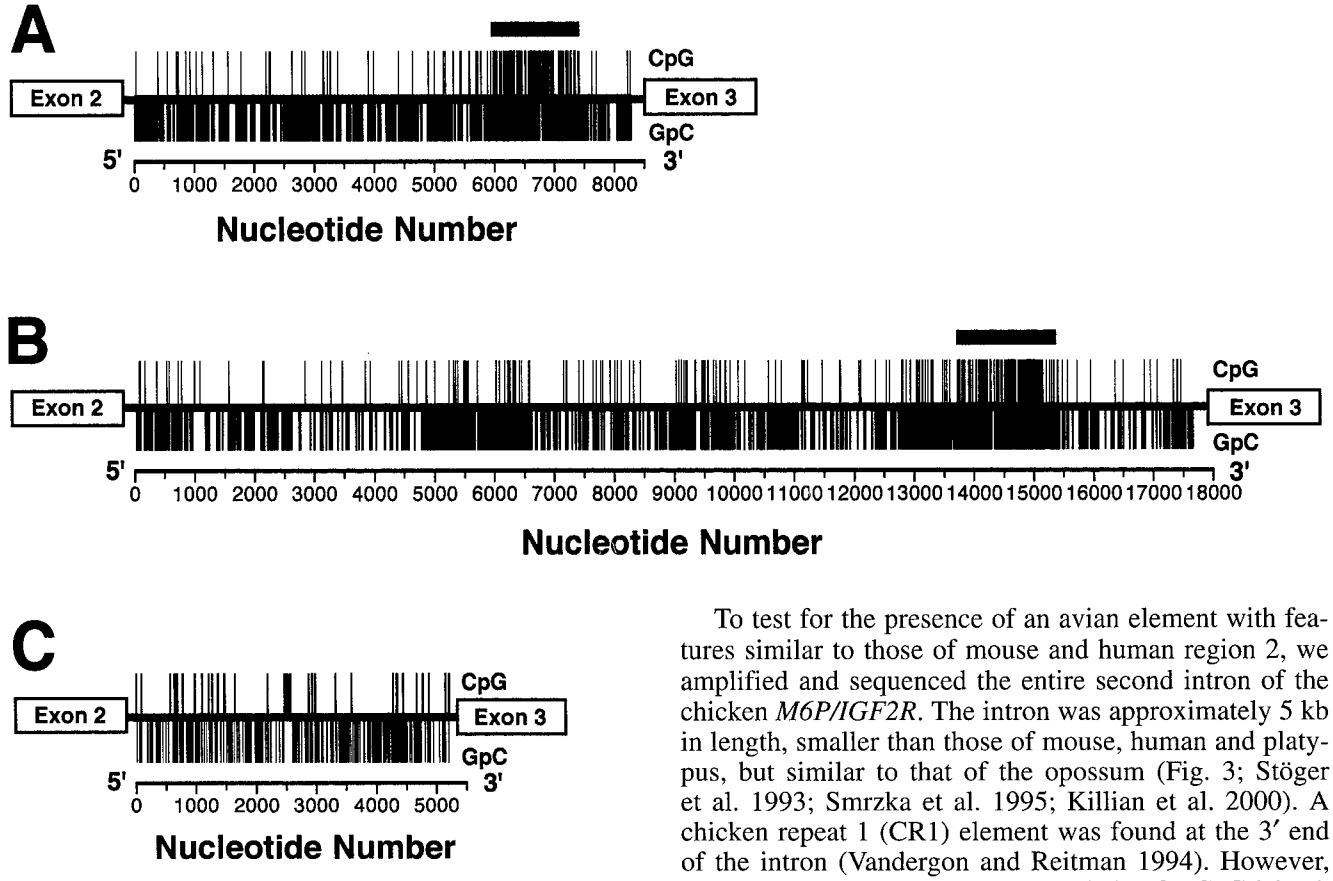

Fig. 3A–C Distribution of CpG and GpC dinucleotides in mouse, human and chicken *M6P/IGF2R* intron 2. **A** The positioning of CpG and GpC sites in the mouse *M6P/IGF2R* intron 2 reveals a 1.5-kb CpG island (Wutz et al. 1997) indicated by a *solid horizontal bar*. **B** The human *M6P/IGF2R* intron 2 also contains a CpG island (Smrzka et al. 1995). **C** Chicken intron 2 lacks a CpG island

is currently unclear. Monoallelic expression of the *M6P/IGF2R* in mice is purported to depend on differential methylation of a CpG island in the second intron (region 2; Stöger et al. 1993). Hybridization studies initially indicated that region 2 was present in primates, mice and birds (Stöger et al. 1993). Although a homologous element was reported in humans (Smrzka et al. 1995), the human *M6P/IGF2R* is expressed from both alleles (Kalscheuer et al. 1993; Ogawa et al. 1993). The relationship of mouse region 2 to that of the human is unclear, as there is no significant sequence homology between them. Nevertheless, they share a high level of organizational similarity with human region 2, also in intron 2, containing a CpG island with maternal-specific methylation. In contrast to initial indications that region 2 was widely conserved (Stöger et al. 1993; Smrzka et al. 1995), the platypus and opossum *M6P/IGF2R* do not have a region 2-like element (Killian et al. 2000). Since the opossum *M6P/IGF2R* is expressed only from the maternally inherited allele, either the mechanism of *M6P/IGF2R* imprinting in the opossum differs from that of the mouse, or region 2 may not be the primary regulator of imprinting at this locus.

To test for the presence of an avian element with features similar to those of mouse and human region 2, we amplified and sequenced the entire second intron of the chicken *M6P/IGF2R*. The intron was approximately 5 kb in length, smaller than those of mouse, human and platypus, but similar to that of the opossum (Fig. 3; Stöger et al. 1993; Smrzka et al. 1995; Killian et al. 2000). A chicken repeat 1 (CR1) element was found at the 3′ end of the intron (Vandergon and Reitman 1994). However, the intron had no features characteristic of a CpG island. It did not have a high G+C content or any clustering of CpG dinucleotides (Fig. 3), and G/C boxes, a regular occurrence in CpG islands (Gardiner-Garden and Frommer 1987), are entirely missing from the sequence. Thus, although region 2 was reported to be present in birds (Stöger et al. 1993), our results demonstrate that chicken *M6P/IGF2R* intron 2 does not contain a CpG island analogous to those in mice and humans (Stöger et al. 1993; Smrzka et al. 1995). The absence of region 2 in the chicken, platypus and opossum, together with the lack of sequence homology between mouse region 2 and human region 2, suggest that such elements may have evolved independently in several mammalian orders. Analysis of the *M6P/IGF2R* in additional marsupial and eutherian mammals is necessary to clarify the involvement of region 2 in regulating imprinting of the *M6P/IGF2R*.

In conclusion, our demonstration that chicken *IGF2* and *M6P/IGF2R* are both expressed biallelically is consistent with the predictions of the genetic conflict hypothesis of imprint evolution (Moore and Haig 1991). Our data also support the idea that the reciprocal imprinting of these genes in viviparous mammals is closely related to their opposing effects on embryonic growth.

Acknowledgements We would like to acknowledge the expert technical assistance of Michael Kulik and the expertise of the Duke DNA sequencing facility and thank the members of the Jirtle laboratory for critical reading of the manuscript. Laboratory animal technician support of the NCSU Scott Hall poultry facility is greatly appreciated. This study was supported by NIH grants CA25951 and ES08823 and DOD grant DAMD17-98-1-8305.

References

Abdel-Hameed F, Shoffner RN (1971) Intersexes and sex determination in chickens. Science 172:962–964

Barlow DP, Stöger R, Herrmann BG, Saito K, Schweifer N (1991) The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. Nature 349:84–87

Clairmont KB, Czech MP (1989) Chicken and *Xenopus* mannose 6-phosphate receptors fail to bind insulin-like growth factor II. J Biol Chem 264:16390–16392

Dahms NM, Brzycki-Wessell MA, Ramanujam KS, Seetharam B (1993) Characterization of mannose 6-phosphate receptors (MPRs) from opossum liver: opossum cation-independent MPR binds insulin-like growth factor-II. Endocrinology 133: 440–446

DeChiara TM, Efstratiadis A, Robertson EJ (1990) A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. Nature 345:78–80

Feil R, Khosla S, Cappai P, Loi P (1998) Genomic imprinting in ruminants: allele-specific gene expression in parthenogenetic sheep. Mamm Genome 9:831–834

Gardiner-Garden M, Frommer M (1987) CpG islands in vertebrate genomes. J Mol Biol 196:261–282

Haig D, Graham C (1991) Genomic imprinting and the strange case of the insulin-like growth factor II receptor. Cell 64:1045–1046

Hu JF, Vu TH, Hoffman AR (1995) Differential biallelic activation of three insulin-like growth factor II promoters in the mouse central nervous system. Mol Endocrinol 9:628–636

Kalscheuer VM, Mariman EC, Schepens MT, Rehder H, Ropers H-H (1993) The insulin-like growth factor type-2 receptor gene is imprinted in the mouse but not in humans. Nat Genet 5:74–78

Killian JK, Byrd JC, Jirtle JV, Munday BL, Stoskopf MK, MacDonald RG, Jirtle RL (2000) *M6P/IGF2R* imprinting evolution in mammals. Mol Cell 5:707–716

Kornfeld S (1992) Structure and function of the mannose 6-phosphate/insulinlike growth factor II receptors. Annu Rev Biochem 61:307–330

Koski LB, Sasaki E, Roberts RD, Gibson J, Etches RJ (2000) Monoalleleic transcription of the insulin-like growth factor-II gene (*Igf2*) in chick embryos. Mol Reprod Dev 56:345–352

Lau MM, Stewart CE, Liu Z, Bhatt H, Rotwein P, Stewart CL (1994) Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. Genes Dev 8:2953–2963

Leighton PA, Ingram RS, Eggenschwiler J, Efstratiadis A, Tilghman SM (1995) Disruption of imprinting caused by deletion of the *H19* gene region in mice. Nature 375:34–39

Ludwig T, Eggenschwiler J, Fisher P, D'Ercole AJ, Davenport ML, Efstratiadis A (1996) Mouse mutants lacking the type 2 *IGF receptor (IGF2R)* are rescued from perinatal lethality in *Igf2* and *Igf1r* null backgrounds. Dev Biol 177:517–535

Moore T, Haig D (1991) Genomic imprinting in mammalian development: a parental tug-of-war. Trends Genet 7:45–49

Nolan CM, Lawlor MA (1999) Variable accumulation of insulin-like growth factor II in mouse tissues deficient in insulin-like growth factor II receptor. Int J Biochem Cell Biol 31:1421–1433

Nolan CM, Kyle JW, Watanabe H, Sly WS (1990) Binding of insulin-like growth factor II (IGF-II) by human cation- independent mannose 6-phosphate receptor/IGF-II receptor expressed in receptor-deficient mouse L cells. Cell Regul 1:197–213

Ogawa O, McNoe LA, Eccles MR, Morison IM, Reeve AE (1993) Human insulin-like growth factor type I and type II receptors are not imprinted. Hum Mol Genet 2:2163–2165

Ohno S, Kittrell WA, Christian LC, Stenius C, Witt GA (1963) An adult triploid chicken (*Gallus domesticus*) with a left ovotestis. Cytogenetics 2:42–49

O'Neill MJ, Ingram RS, Vrana PB, Tilghman SM (2000) Allelic expression of *IGF2* in marsupials and birds. Dev Genes Evol 210:18–20

Overall M, Bakker M, Spencer J, Parker N, Smith P, Dziadek M (1997) Genomic imprinting in the rat: linkage of *Igf2* and *H19* genes and opposite parental allele-specific expression during embryogenesis. Genomics 45:416–420

Pham NV, Nguyen MT, Hu JF, Vu TH, Hoffman AR (1998) Dissociation of *IGF2* and *H19* imprinting in human brain. Brain Res 810:1–8

Rainier S, Johnson LA, Dobry CJ, Ping AJ, Grundy PE, Feinberg AP (1993) Relaxation of imprinted genes in human cancer. Nature 362:747–749

Sarvela P (1973) Adult parthenogenetic chickens. Nature 243:171

Smrzka OW, Fae I, Stöger R, Kruzbauer R, Fischer GF, Henn T, Weith A, Barlow DP (1995) Conservation of a maternal-specific methylation signal at the human IGF2R locus. Hum Mol Genet 4:1945–1952

Stöger R, Kubicka P, Liu CG, Kafri T, Razin A, Cedar H, Barlow DP (1993) Maternal-specific methylation of the imprinted mouse *Igf2r* locus identifies the expressed locus as carrying the imprinting signal. Cell 73:61–71

Vandergon TL, Reitman M (1994) Evolution of chicken repeat 1 (CR1) elements: evidence for ancient subfamilies and multiple progenitors. Mol Biol Evol 11:886–898

Vu TH, Hoffman AR (1994) Promoter-specific imprinting of the human insulin-like growth factor-II gene. Nature 371:714–717

Wang ZQ, Fung MR, Barlow DP, Wagner EF (1994) Regulation of embryonic growth and lysosomal targeting by the imprinted *Igf2/Mpr* gene. Nature 372:464–467

Wutz A, Smrzka OW, Schweifer N, Schellander K, Wagner EF, Barlow DP (1997) Imprinted expression of the *Igf2r* gene depends on an intronic CpG island. Nature 389:745–749

Yandell CA, Dunbar AJ, Wheldrake JF, Upton Z (1999) The kangaroo cation-independent mannose 6-phosphate receptor binds insulin-like growth factor II with low affinity. J Biol Chem 274:27076–27082

Zhou M, Ma Z, Sly WS (1995) Cloning and expression of the cDNA of chicken cation-independent mannose-6-phosphate receptor. Proc Natl Acad Sci USA 92:9762–9766

Dev Genes Evol (2001) 211:184–189
DOI 10.1007/s004270100133

SHORT COMMUNICATION

Akie Nakayama · Yutaka Satou · Nori Satoh

Isolation and characterization of genes that are expressed during *Ciona intestinalis* metamorphosis

Received: 6 November 2000 / Accepted: 5 December 2000 / Published online: 14 March 2001

Abstract In ascidians, the events of metamorphosis transform the non-feeding, mobile tadpole larva into a filter-feeding, fixed juvenile, and the process involves rearrangements of cells, two organs and physiological changes. Differential screening was used to isolate two genes that are not expressed in swimming larvae but are expressed immediately after the initiation of metamorphosis in *Ciona intestinalis*. One of the genes, *Ci-meta1*, encodes a polypeptide with a putative secretion signal sequence, 6 epidermal growth factor (EGF)-like repeats and 13 calcium-binding EGF-like repeats. The gene begins to be expressed immediately after the beginning of metamorphosis in the adhesive organ and is likely to be associated with the signal response for metamorphosis. Another gene named *Ci-meta2* encodes a protein with a putative secretion signal and three thrombospondin type-1 repeats. *Ci-meta2* gene expression begins at the larval stage and is upregulated in the metamorphosing juveniles. *Ci-meta2* expression is found in three regions; the adhesive organ which is also associated with settlement, the neck region between the trunk and the tail of the larva which is associated with tail resorption, and dorsal regions of the trunk which correspond to the location of the siphon primordium. This gene may be involved in the dynamic arrangement of cells during ascidian metamorphosis.

Keywords Ascidians · Metamorphosis · Genes · Cell signaling · Cell rearrangement

Edited by R.P. Elinson

A. Nakayama (✉) · Y. Satou · N. Satoh
Department of Zoology, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan
e-mail: akie@ascidian.zool.kyoto-u.ac.jp
Tel.: +81-75-7534095, Fax: +81-75-7051113

Introduction

The events of ascidian metamorphosis transform the non-feeding, mobile tadpole larva into a filter-feeding, fixed juvenile (Cloney 1982; Satoh 1994; Jeffery and Swalla 1997). Metamorphosis involves numerous rapid morphogenetic movements and physiological changes that are initiated at the moment of settlement. According to Cloney (1978, 1982), the tadpole organs and tissues can be categorized into three groups: transitory larval organs (TLO), prospective juvenile organs (PJO), and larval-juvenile organs or tissues (LJO). TLO function in larval locomotion (notochord and muscle), sensory input (otolith and ocellus), and settlement (papillae). They are fully differentiated in the larva, but are destroyed or, in the case of the fins, lost during metamorphosis. PJO are in an arrested state of development in the larva, and they may become functional either shortly after settlement or following further histogenesis. LJO are organs or tissues that function in both the larval and post-larval phases of the life cycle. Therefore, metamorphosis is accomplished by dynamic changes in all three types of tissues and organs. In addition, the beginning of metamorphosis can be triggered in the non-feeding larvae by artificial cues (Satoh 1994; Degnan et al.1997), and processes of metamorphosis are highly synchronous among individuals. The process of metamorphosis takes place within a very short period of time; the tail resorption of *Ciona*, for example, is accomplished in 30 min.

Recently, Lavin and Degnan and their colleagues have begun to characterize molecules from ascidian tadpole larvae that are expressed in response to specific environmental cues that trigger metamorphosis. They identified a novel gene (*Hemps*), which encodes a protein with a putative secretion signal sequence and four epidermal growth factor (EGF)-like repeats, which is a key regulator of metamorphosis in the ascidian *Herdmania curvata* (Arnold et al. 1997). Expression of *Hemps* increases markedly when the swimming tadpole larva becomes competent to undergo metamorphosis and then during the first 24 h of metamorphosis (Eri et al. 1999). The

Mammalian Genome 12, 513–517 (2001).
DOI: 10.1007/s003350020026

Mammalian Genome
Incorporating Mouse Genome

Marsupials and Eutherians reunited: genetic evidence for the Theria hypothesis of mammalian evolution

J. Keith Killian,[1] Thomas R. Buckley,[2] Niall Stewart,[3] Barry L. Munday,[3] Randy L. Jirtle[1]

[1]Departments of Radiation Oncology and Pathology, Duke University Medical Center, Box 3433, Durham, North Carolina 27710, USA
[2]Department of Biology, Duke University, Durham, North Carolina, 27708, USA
[3]School of Biomedical Science, University of Tasmania, Launceston, Tasmania 7250, Australia

Received: 8 December 2000 / Accepted: 01 February 2001

Abstract. The three living monophyletic divisions of Class Mammalia are the Prototheria (monotremes), Metatheria (marsupials), and Eutheria ('placental' mammals). Determining the sister relationships among these three groups is the most fundamental question in mammalian evolution. Phylogenetic comparison of these mammals by either anatomy or mitochondrial DNA has resulted in two conflicting hypotheses, Theria and Marsupionta, and has fueled a "genes versus morphology" controversy. We have cloned and analyzed a large nuclear gene, the mannose 6-phosphate/insulin-like growth factor II receptor (*M6P/IGF2R*), from representatives of all three mammalian groups, including platypus, echidna, opossum, wallaby, hedgehog, mouse, rat, rabbit, cow, pig, bat, tree shrew, colugo, ringtail lemur, and human. Statistical analysis of this nuclear gene unambiguously supports the morphology-based Theria hypothesis that excludes monotremes from a clade of marsupials and eutherians. The *M6P/IGF2R* was also able to resolve the finer structure of the eutherian mammalian family tree. In particular, our analyses support sister group relationships between lagomorphs and rodents, and between the primates and Dermoptera. Statistical support for the grouping of the hedgehog with Feruungulata and Chiroptera was also strong.

Introduction

The Class Mammalia includes the egg-laying monotremes, as well as the more familiar live-birthing members. The viviparous mammals are further divided into marsupials and eutherians, depending on reproductive parameters such as placentation and the extent of intrauterine gestation of the developing offspring. Marsupial young generally exit the womb at a significantly premature stage of development relative to eutherians (Griffiths 1999; Renfree and Shaw 1999). A sister relationship between the marsupials and eutherians, exclusive of monotremes, was deduced from anatomical data (Marshall 1979). Morphological synapomorphies that distinguish marsupials and eutherians from monotremes include: olfactory bulb mitral cell organization, braincase architecture, cranial nerve distribution, inner ear cochlear architecture, mammary glands with teats, tooth enamel, molar dentition, and foot anatomy (Lewis 1983; Marshall 1979).

Recent analyses of tribosphenic molar fossils concluded that monotremes and marsupials/eutherians descended from ancestors in geographically distinct areas: living monotremes are descendants of an ancient mammalian clade endemic to Gondwanan land masses, whereas marsupials and eutherians shared a common ancestor upon Laurasian continents (Luo et al. 2001). In accord with the conflict hypothesis for the origin of genomic imprinting (Moore and Haig 1991), the differential expression of parental alleles of certain fetal growth-regulatory genes is another therian apomorphy shown to not be present in Prototherian ancestors (Killian et al. 2000). Furthermore, delineation of retroposon elements in the three mammalian groups supports the earlier divergence of Prototheria (Gilbert and Labuda 2000).

The value and accuracy of decades of morphological study have been discounted recently by mitochondrial DNA inference, which has reinvigorated Gregory's claim that monotremes are highly-derived marsupials (Gregory 1947; Janke et al. 1997; Penny et al. 1999). The accurate resolution of the mammalian family tree is essential to determining whether the unique physical attributes of marsupials and eutherians, as well as the imprinting of certain genes, have evolved once or numerous times. If the Marsupionta hypothesis proves correct, then either these characteristics represent phenotypic and epigenetic convergences in marsupials and eutherians, or else they were gained and then lost in the monotreme lineage. Substantiation of the Theria hypothesis would implicate a more parsimonious single origin for these features.

The incongruent mammalian family trees resulting from mitochondrial sequence analysis and morphological characters have contributed to a divisive debate over their utilities in phylogenetic inference (Gura 2000). Thus far, only small nuclear genes have been applied to the Theria/Marsupionta question (i.e., <1.5 kb open reading frame; Kullander et al. 1997; Toyosawa et al. 1999), and they have failed to convincingly validate either hypothesis. In an effort to resolve the higher order structure of the mammalian family tree, we have performed phylogenetic inference based on the mannose-6-phosphate/insulin-like growth factor II receptor (*M6P/IGF2R*) gene from representatives of all three mammalian groups, using chicken and fish homologs as outgroups.

The 48 exons of *M6P/IGF2R* produce a large 9-kb transcript, which is translated into a protein of over 2400 amino acids (Jirtle 1999). Studies in mice have established that the *M6P/IGF2R* is an essential, single-copy nuclear gene that is critically involved in embryogenesis (Jirtle 1999). The *M6P/IGF2R* has a long and well-established evolutionary history in the animal kingdom, with definitive homologs identified in fish, aves, and mammals (Jirtle 1999; Killian et al. 2000; Nadimpalli et al. 1999). Biochemical evidence also indicates the presence of a homologous gene in invertebrates (Lakshmi et al., 1999). We have isolated and sequenced novel *M6P/IGF2R* homologs from monotremes, marsupials, and eutherians, including echidna, wallaby, hedgehog, pig, bat, tree shrew, colugo, and ringtail lemur. To generate the mammalian phylogeny, these sequences were used in addition to previously characterized *M6P/IGF2R* homologs from chicken, platypus, opossum, mouse, rat, cow, and human (Jirtle 1999; Killian et

The M6P/IGF2R nucleotide sequence data reported in this paper have been submitted to GenBank and have been assigned the accession numbers: AF339157-163 (bat, rabbit, wallaby, lemur, tree shrew, hedgehog, colugo), AF339885 (pig), AF342813 (opossum), and AF342814 (echidna).

Correspondence to: Randy L. Jirtle; E-mail: jirtle@radonc.duke.edu

Table 1. Cross-species *M6P/IGF2R* PCR primers.

Forward Primers 5′ to 3′		Reverse primers 5′ to 3′	
F primer[a]	Sequence	R primer[a]	Sequence
311F	CTGTGCAGTTACACATGGGAAGC	1078R	GGCATACTCAGTGATCCACTC
589F	GGAACTCCTGAATTTGTAACTGCCACAG	1498R	GTAGGTGCAGTCCACCTC
617F	TGTGTGCATTACTTTGAATGGAGGAC	1505R	GTGAAGAAGTAGGTGCAGTC
1428F	GGCTTTCAGCGGATG	1642R	CCATCCACTGCTTCCCA
1450F	ATGAGTGTCATAAACTTTGAGTGC	1921R	GCACTTTCTAGATCACCTGG
1451F	TGTCATAAACTTTGAGTGCAA	2043R	TCAAAAGAAAACCCTGCCTG
1650F	GAACAGAATTGGGAAGC	4225R	GTCATTGAAAGGTGGGCAGGC
1655F	CAGAATCAGAACAGAATTGGGAAGCTGTGG	5153R	AAGCCTCATAACCACCAGTGCG
1657F	TTGGGAAGCAGTGGATG	6018R	CTCCCAATGTCCTCATCTTCATCAC
1929F	CCAGGTGATCTAGAAAGTGC	6534R	CCATTCACCATCTGCACCTCCTG
2410F	TACCTACAACTTCCGGTGG	6877R	GGGTCACAGTGGAAAAAGAT
2415F	AACTCTACCTACAACTTCCGGTGGTACAC	7250R	ATCTCTTCCATCAGCCACTC
2883F	CAACTGGGAGTGTGTGGTC	7252R	CTGGATTTCTTCCATAAGCCA
4230F	TTTGAGTGGCGAACCCAGTATGCCTG	7588R	GTCCTCGTCGCTGTCGTCATG
5834F	GTGAATGGTGATCGTTGCCCTCCAG	oligo 8	GACCACGCGTATCGATGTCGACT$_{16}$V
5886F	CCCTTCATATTCAATGGGAAGAGC	oligo 9	GACCACGCGTATCGATGTCGAC
6129F	CCTCCAAAGAAGATGGAGTGC		
7263F	GAGAATGAAACGGAGTGGCTTATGGA		

[a] Oligonucleotides are numbered relative to the 9090-bp human M6P/IGF2R transcript (GenBank Accession number: NM_000876). PCR primer use is described in Materials and methods. Oligo 8 and oligo 9 are commercially available (Roche Boehringer Mannheim, Indianapolis, Ind.).

al. 2000). Inclusion of these novel sequences allowed us to split long branches in the phylogeny and to test the ability of the *M6P/IGF2R* to resolve the interordinal relationships among mammals.

Materials and methods

Tissue samples. Short-beaked echidna (*Tachyglossus aculeatus*) samples were obtained from animal victims of car accidents. Samples of the Tasmanian subspecies of *Macropus rufogriseus,* the Bennett's wallaby, were obtained from Lenah Game Meats, Launceston, Tasmania. European hedgehog (*Erinaceus europaeus*) tissue was obtained postmortem from an ailing animal euthanized by a veterinarian at Duke University Medical Center Vivarium. New Zealand white rabbit tissue was obtained from the Duke University Medical Center Vivarium. Domestic pig (*Sus scrofa*) muscle was obtained from Neese's Sausage, Burlington, North Carolina, USA. Brown bat (*Myotis lucifugus*) samples were provided by the North Carolina Wildlife Commission. Tree shrew (*Tupaia glis*) tissue was donated from the archives of David Fitzpatrik, Duke University Medical Center. Philippine flying lemur (*Cynocephalus volans*) tissue was a kind gift from the Field Museum, Chicago, USA. Ringtail lemur (*Lemur catta*) tissues were from the archives of the Duke University Primate Facility. In most cases, tissues were dispatched to Duke University Medical Center in RNALater (Ambion, Austin, TX) until DNA and RNA extraction.

M6P/IGF2R cloning. Total RNA was isolated from 50 mg of animal tissue by homogenization in RNA-Stat 60 (Tel-Test, Friendswood, Texas). First-strand cDNA was synthesized from 1–5 μg of total RNA by using SuperScript II (Life Technologies, Baltimore, Md.) and oligo 8 (Roche Boehringer Mannheim, Indianapolis, Ind.) as the primer for reverse transcription (Table 1). We have previously described the design and implementation of a battery of non-degenerate PCR primers for cross-species *M6P/IGF2R* amplification of opossum and platypus orthologes (Killian et al. 2000). Briefly, the use of primers 311F, 589F, and/or 617F with 1078R yielded a correct amplimer for all species tested and allowed us to obtain species-specific *M6P/IGF2R* sequence (Table 1). To reduce the complexity of the PCR template, we then performed 3′-RACE according to the manufacturer's protocol (Life Technologies) with a species-specific *M6P/IGF2R* primer and oligo 9 (Table 1) (Roche Molecular Biochemicals, Indianapolis, Ind.). Typical PCR reactions used 1% of the RNA-to-cDNA RT products, 1.5 U Expand Long Template DNA polymerase mix (Roche Molecular Biochemicals) 1 μM of each primer, and 500 μM dNTPs in a 50 μl PCR reaction volume (94°C × 20 s, 55°C × 5 s, and 68°C × 3 min for 35 cycles). Unpurified 3′-RACE product (1 μl) was then used as template in additional cross-species *M6P/IGF2R* PCR amplifications by using various combinations of non-degenerate cross-species *M6P/IGF2R* oligos and identified species-specific oligos; the PCR reaction parameters were identical to those previously described. PCR products were analyzed by agarose gel electrophoresis, and appropriately sized fragments were excised, purified (GenElute, Sigma Chemical Co., St. Louis, Mo.), and PCR sequenced. A partial fish (*Xiphophorus maculatus* × *Xiphophorus helleri*) *M6P/IGF2R* sequence was obtained from GenBank (Accession Number AJ278449).

Phylogenetic methods. M6P/IGF2R amino acid sequences were aligned using ClustalX (Higgins et al. 1992), with regions of ambiguous homology excluded prior to phylogenetic analysis. The assumption of amino acid frequency stationarity among taxa (i.e., constant amino acid frequencies among lineages) was evaluated using χ^2 tests in Tree-Puzzle 5.0 (Strimmer and von Haeseler 1996). Phylogenetic trees were constructed by using maximum parsimony (MP), maximum likelihood (ML), minimum evolution (ME), and split decomposition (reviewed by Swofford et al. 1996). We performed the MP analyses in PAUP*4.0b4a (Swofford 1998), using equal weights and the Blosum 80 step matrix (Henikoff and Henikoff 1992). All MP analyses employed a heuristic search strategy with start trees obtained via stepwise addition with 10 random addition replicates followed by TBR branch swapping. ML trees were constructed with a global search by using Proml from the Phylip 3.6a package (Felsenstein 1993) under the Dayhoff et al. (1978) model with rate categories constrained to fit a discrete gamma distribution (Yang 1994). The gamma distribution was estimated by using Tree-Puzzle 5.0. ME trees were constructed in PAUP*4.0b4a from distances estimated in Protdist from Phylip 3.6a under the Dayhoff et al. (1978) model with a gamma correction as described above. Log Determinant (Lockhart et al. 1994) distances with conserved site removal were estimated in SplitsTree 3.0 (Huson 1998). Shimodaira-Hasegawa (1999) tests were performed with PAML 3.0 (Yang 1997) under the Dayhoff et al. (1978) and JTT-F (Jones et al. 1992) models with among-site rate variation described by using a discrete gamma distribution. The test statistic for the Shimodaira-Hasegawa (1999) test was obtained by comparing the ML topology with the most likely topology under the constraint that the marsupial and monotreme taxa form a monophyletic group.

Results and discussion

Alignment of the *M6P/IGF2R* genes yielded a data set containing 2257 amino acid sites from each of the 15 mammals and the chicken outgroup. Of these sites, 1408 were varied and 973 were parsimony informative. We report herein only the analyses of the amino acid sequences that contained sufficient variation to generate robust estimates of the mammalian phylogeny. All of our phylogenetic analyses strongly support the grouping of the eutherian mammals with the marsupials, consistent with the traditional. Theria hypothesis, but not with the Marsupionta hypothesis (Fig. 1). In particular, MP, ML, and ME analyses of *M6P/IGF2R* sequences resulted in 100% bootstrap support for the grouping of marsupials

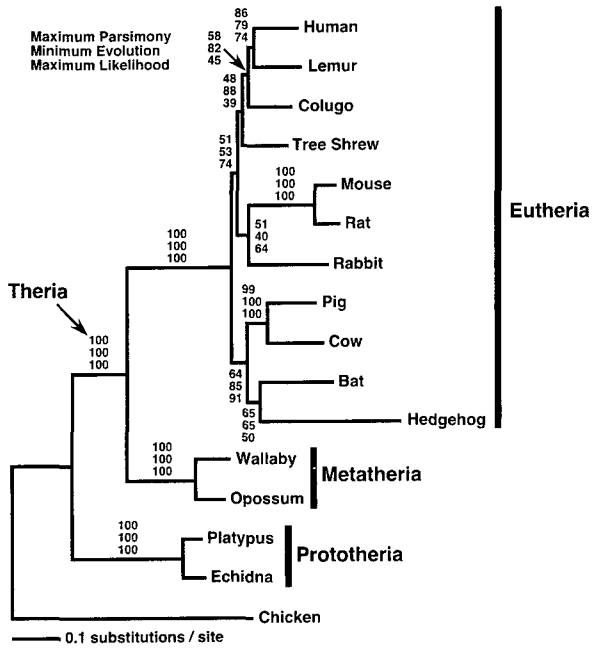

Fig. 1. Phylogenetic tree constructed by maximum likelihood from *M6P/IGF2R* amino acid sequences. Bootstrap values showing statistical support are given adjacent to each node. From the top: weighed maximum parsimony, minimum evolution [Dayhoff et al. (1978) model with a Γ-correction], and maximum likelihood [Dayhoff et al. (1978) model with distributed rates].

and eutherian mammals (Fig. 1). Additionally, using the conservative Shimodaira-Hasegawa (1999) test, we were able to reject ($P < 0.001$) the null hypothesis that there is no difference in likelihood between the ML topology and the constrained Marsupionta tree. Therefore, the results of this test indicate that the data are inconsistent with the Marsupionta hypothesis.

Because the χ^2 tests for amino acid frequency stationarity indicated that the marsupials and the chicken had deviating amino acid compositions, we constructed minimum evolution trees and split decomposition networks from log Determinant distances (Lockhart et al. 1994) with conserved site removal (Penny et al. 1999; Waddell et al. 1999a). These analyses (data not shown) still supported the Theria clade with 97% bootstrap support under split decomposition. In addition, to test the sensitivity of our analyses to the outgroup selected, we restricted our alignment to 764 amino acid sites in order to use a partial fish *M6P/IGF2R* sequence. The addition of the fish sequence had no effect on the relationships among the three major mammalian lineages, with the Theria hypothesis receiving 97% bootstrap support under weighted maximum parsimony, and 87% bootstrap support under minimum evolution. This observation indicates that our results are insensitive to outgroup selection. Therefore, support for the Theria hypothesis is strongly independent of the method of phylogenetic reconstruction employed and does not stem from an incorrect rooting of the mammalian tree.

Our phylogenetic analyses also support (64% to 91%) the grouping of the hedgehog with the Feruungulata. This relationship was predicted by Waddell et al. (1999b), but contradicts that of some analyses of whole mitochondrial genomes (Mouchaty et al. 1999). We also note that our grouping of hedgehog with the Feruungulata is consistent with other nuclear gene analyses (Stanhope et al. 1998) and with maximum likelihood analyses of 1st and 2nd codon positions from whole mitochondrial genomes (Sullivan and Swofford 1997). The estimates of bootstrap support are moderate (50%–65%) for the grouping of the hedgehog with the bat sequence. This relationship is not surprising, given that Mouchaty et al. (2000) observed a sister group relationship between another insectivore, the mole, and the bat. The observation that phylogenetic analyses of a number of nuclear genes support the affinity of the hedgehog with the Feruungulata and/or Chiroptera (Stanhope et al. 1998; Springer et al. 1999; Waddell et al. 1999b; this study) indicates that the hedgehog is almost certainly misplaced in most studies of whole mitochondrial genomes.

All of our analyses support the Eucharonta clade (Waddell et al. 1999b), containing the orders Primates, Dermoptera, and Scandentia. Estimates of bootstrap support are low under MP (48%) and ML (39%), yet they are high under ME (88%). This grouping appears to conflict with the analyses of whole mitochondrial genomes described by Schmitz et al. (2000); however, the bootstrap support values they obtained for a Lagomorph + Scandentia clade are low, and a Scandentia + Primate grouping could not be rejected when the complete data set was used. The phylogenetic analyses of four nuclear genes and three mitochondrial genes described by Teeling et al. (2000) also yielded a Dermoptera + Primate clade with moderate bootstrap support, consistent with the results of our study.

Interestingly, all of our phylogenetic analyses support the grouping of rodents and lagomorphs in the Glires clade. The Glires grouping is somewhat controversial, although support from morphological (Liu and Miyamoto 1999), paleontological (Meng et al. 1994), and embryological (Lucket and Hartenberger 1993) studies is strong. Estimates of statistical support from various molecular data sets (Halanych 1998; Huchon et al. 1999; Waddell et al. 1999a; Robinson-Rechavi et al. 2000) have typically been moderate, as is the bootstrap estimate we have obtained under ML (64%). We suspect that the generally low level of support for the Glires clade observed in all molecular studies to date stems in part from poor taxon sampling of both lagomorphs and rodents.

The phylogenetic relationships among Eutherian mammals that we have recovered are largely consistent with the predictions of Waddell et al. (1999b). These observations are important because they indicate that the *M6P/IGF2R* gene provides a good source of characters for resolving the higher-order mammalian phylogeny. Additionally, these results indicate that the *M6P/IGF2R* gene is a promising candidate for further resolving relationships among other vertebrate groups.

The strong genetic support that we have obtained for the Theria hypothesis strikingly contradicts the mammalian family tree supported by statistical analyses of whole mitochondrial genomes (Janke et al. 1996; Penny and Hasegawa 1997). Interestingly, Janke et al. (1997) were able to reject the Theria hypothesis only by using the *Xenopus laevis* mitochondrial sequence as an outgroup, and not with the much more appropriate chicken sequence. The statistical tests of topology (Kishino and Hasegawa 1989), implemented by Janke et al. (1997), also used substitution models that assumed equal substitution rates among sites. This assumption is known to seriously bias tests of phylogenetic hypotheses in many situations (e.g., Sullivan and Swofford 1997). Waddell et al. (1999a) observed that support for the Marsupionta hypothesis was very poor when conserved mitochondrial tRNA sequences were analyzed, especially under ML. Furthermore, Penny et al. (1999) determined with the use of the distance Hadamard method that there was strong conflict from within the mitochondrial data for grouping monotremes with the marsupials.

Our findings also contradict the DNA hybridization studies of Kirsch and Mayer (1998). Their results supported the Marsupionta hypothesis; however, they cautioned that they could be biased by shifts in GC nucleotide content over the tree. There is also debate concerning the utility of DNA hybridization data for divergences older than 50 million years (Hillis et al. 1996). Thus, the relationships among the three major mammalian lineages appear not to be definitely resolved by either the mitochondrial or DNA hybridization studies.

The results presented herein provide the first genetic sequence data from a large nuclear gene that reconcile genes, epigenetics,

and morphology in understanding the mammalian family tree, and they unambiguously support the accuracy of the Theria hypothesis. In the Triassic period, monotreme ancestors diverged from the mammalian ancestors that in the late Jurassic/early Cretaceous periods diverged to give rise to marsupial and eutherian mammals. It is significant that apomorphies of the therian ancestors, such as the braincase, cranial nerve architecture, and reproductive physiology do not need to be reclassified as convergences, a problem created previously by both mitochondrial DNA and DNA hybridization studies. Furthermore, our findings support the postulate that the imprinting of growth-regulatory genes, such as *M6P/IGF2R* is an apomorphy of viviparous mammals, further distinguishing them from oviparous taxa.

Acknowledgments. The authors thank the following for kindly providing tissues: Steven Atkins, Lenah Game Meats, Animals Exotique, Neese's Sausage, Wild Kingdom Animal Removal, David Fitzpatrik, The Field Museum, and the Duke University Primate Center. We further wish to thank the Duke University DNA analysis facility for DNA sequencing, and Kay Nolan for critical review of the manuscript. This study was supported by National Institutes of Health grants CA25951 and ES08823, Department of Defense grant DAMD17-98-1-8305, Sumitomo Chemical Company, Ltd., and AstraZeneca Pharmaceuticals, Ltd. T.R. Buckley thanks Clifford Cunningham and the Duke University Cancer Center for support.

References

Dayhoff M, Schwartz R, Orcutt B (1978) A model of evolutionary change in proteins. In *Atlas of Protein Sequence and Structure,* Dayhoff MO, ed. (Washington, DC: National Biomedical Research Foundation), pp 345–352

Felsenstein J (1993) PHYLIP: phylogeny inference package, version 3.6a, Department of Genetics, Univ. of Washington, Seattle

Gilbert N, Labuda D (2000) Evolutionary inventions and continuity of CORE-SINEs in mammals. J Mol Biol 298, 365–377

Gregory WK (1947) The monotremes and the palimpsest theory. Bull Am Mus Nat Hist 88, 1–52

Griffiths M (1999) Monotremes. In *Encyclopedia of Reproduction,* Knobil E, Neill JD, eds. (San Diego, Calif.: Academic Press), pp 295–302

Gura T (2000) Bones, molecules . . . or both? Nature 406, 230–233

Halanych KM (1998) Lagomorphs misplaced by more characters and fewer taxa. Syst Biol 47, 138–146

Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 89, 10915–10919

Higgins D, Bleasby A, Fuchs R (1992) Improved software for multiple sequence alignment. Comput Appl Biosci 8, 189–191

Hillis DM, Mable BK, Moritz C (1996) Applications of molecular systematics. In *Molecular Systematics,* 2nd ed, Hillis DM, Moritz C, Mable BK, eds. (Sunderland, Mass.: Sinauer Associates) pp 515–543

Huchon D, Catzeflis FM, Douzery EJP (1999) Molecular evolution of the nuclear von Willebrand factor gene in mammals and the phylogeny of rodents. Mol Biol Evol 16, 577–589

Huson D (1998) SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics 14, 68–73

Janke A, Gemmell NJ, Feldmaier-Fuchs G, von Haeseler, Paabo S. (1996) The mitochondrial genome of a monotreme. The platypus (Ornithorhynchus anatinus), J Mol Evol 42, 153–159

Janke A, Xu X, Arnason U (1997) The complete mitochondrial genome of the wallaroo (*Macropus robustus*) and the phylogenetic relationship among Monotremata, Marsupialia, and Eutheria. Proc Natl Acad Sci USA 94, 1276–1281

Jirtle RL (1999) Mannose 6-phosphate receptors. In *Encyclopedia of Molecular Biology,* Creidton TE, ed. (New York, NY: Wiley-Liss, Inc.), pp 1441–1447

Jones D, Taylor W, Thornton J (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8, 275–282

Killian JK, Byrd JC, Jirtle JV, Munday BL, Stoskopf MK et al. (2000) *M6P/IGF2R* imprinting evolution in mammals. Mol Cell 5, 707–716

Kirsch JAW, Mayer GC (1998) The platypus is not a rodent: DNA hybridization, amniote phylogeny and the palimpsest theory. Philos Trans R Soc Lond Biol Sci 353, 1221–1237

Kishino H, Hasegawa M (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequences, and the branching order of Hominicidea. J Mol Evol 29, 170–179

Kullander K, Carlson B, Hallbook F (1997) Molecular phylogeny and evolution of the neutrophins from monotremes and marsupials. J Mol Evol 45, 311–321

Lakshmi YU, Radha Y, Hille-Rehfeld A, von Figura K, Kumar NS (1999) Identification of the putative mannose 6-phosphate receptor protein (MPR 300) in the invertebrate unio. Biosci Rep 19, 403–409

Lewis OJ (1983) The evolutionary emergence and refinement of the mammalian pattern of foot architecture. J Anat 137, 21–45

Liu F-G, Miyamoto (1999) Phylogenetic assessment of molecular and morphological data for Eutherian mammals. Syst Biol 48, 54–64

Lockhart PJ, Steel MA, Hendy MD, Penny D (1994) Recovering evolutionary trees under a more realistic model of sequence evolution. Mol Biol Evol 11, 605–612

Luckett WP, Hartenberger J.-L (1993) Monophyly or polyphyly of the order rodencia: possible conflict between morphological and molecular interpretations. J Mammal Evol 1, 127–147

Lucket WP, Hong N (1998) Phylogenetic relationships between the orders Artiodactyl and Cetacea: a combined assessment of morphological and molecular evidence. J Mamm Evol 5, 127–182

Luo Z-X, Cifelli RL, Kielan-Jaworowska Z (2001) Dual origin of tribosphenic mammals. Nature 409, 53–57

Marshall LG (1979) Evolution of metatherian and eutherian (mammalian) characters: a review based on cladistic methodology. Zool J Linn Soc 66, 369–410

Meng J, Wyss AR, Dawson MR, Zhai R (1994) Primitive fossil rodent from Inner Mongolia and its implications for mammalian phylogeny. Nature 370, 134–136

Moore T, Haig D (1991) Genomic imprinting in mammalian development: a parental tug-of-war. Trends Genet 7, 45–49

Mouchatry SK, Gullberg A, Janke A, Arnason U (2000) The phylogenetic position of the Talpidae within the Eutheria based on analysis of complete mitochondrial sequences. Mol Biol Evol 17, 60–67

Nadimpalli SK, Yerramalla UL, Hille-Rehfeld A, von Figura K (1999) Mannose 6-phosphate receptors (MPR 300 and MPR 46) from a teleostean fish (trout). Comp Biochem Physiol B Comp Biochem Mol Biol 123, 261–265

Penny D, Hasegawa M (1997) Molecular systematics. The platypus put in its place. Nature 387, 549–550

Penny D, Hasegawa M, Waddell PJ, Hendy MD (1999) Mammalian evolution: timing and implications from using the log determinant transform for proteins of differing amino acid composition. Syst Biol 48, 76–93

Renfree MB, Shaw G (1999) Marsupials. In *Encyclopedia of Reproduction.* Knobil E, Neill JD, eds (San Diego: Academic Press), pp 104–114

Robinson-Rechavi M, Ponger L, Mouchiroud D (2000) Nuclear gene LCAT supports rodent monphyly. Mol Biol Evol 17, 1410–1412

Schmitz J, Ohme M, Zischler H (2000) The complete mitochondrial genome of *Tupaia belangeri* and the phylogenetic affiliation of Scandentia to other Eutherian orders. Mol Biol Evol 17, 1334–1343

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol 16, 1114–1116

Springer MS, Amrine HM, Burk A, Stanhope MJ (1999) Additional support for Afrotheria and Paenungulata, the performance of mitochondrial versus nuclear genes, and the impact of data partitions with heterogeneous base composition. Syst Biol 48, 65–75

Stanhope MJ, Waddell VG, Madsen O, de Jong W, Hedges SB et al. (1998) Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. Proc Natl Acad Sci USA 95, 9967–9972

Strimmer K, von Haeseler A (1996) Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. Mol Biol Evol 13, 964–969

Sullivan J, Swofford D (1997) Are guinea pigs rodents? The importance of adequate models in molecular phylogenetics. J Mamm Evol 4, 77–86

Swofford D (1998) PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4. (Sunderland, Mass.: Sinauer Associates).

Swofford D, Olsen G, Waddell P, Hillis D (1996) In *Molecular System-*

atics, 2nd ed., Hillis DM, Moritz C, Mable BK, eds. (Sunderland, Mass.: Sinauer Associates), pp 407–514

Teeling EC, Scally M, Kao DJ, Romagnoli ML, Springer MS et al. (2000) Molecular evidence regarding the origin of echolocation and flight in bats. Nature 403, 188–192

Toyosawa S, O'Huigin C, Klein J (1999) The dentin matrix protein 1 gene of prototherian and metatherian mammals. J Mol Evol 48, 160–167

Waddell PJ, Cao Y, Hauf J, Hasegawa M (1999a) Using novel phylogenetic methods to evaluate mammalian mtDNA, including amino acid-invariant sites-logDet plus site stripping, to detect internal conflicts in the data, with special reference to the positions of hedgehog, armadillo, and elephant. Syst Biol 48, 31–53

Waddell P, Okada N, Hasegawa M (1999b) Towards resolving the interordinal relationships of placental mammals. Syst Biol 48, 1–5

Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol 39, 306–314

Yang Z (1997) Phylogenetic analysis by maximum likelihood (PAML). Version 3.0. (London, UK: University College London).

HUMAN MUTATION 18:25–31 (2001)

RAPID COMMUNICATION

Mannose 6-Phosphate/Insulin-Like Growth Factor 2 Receptor (M6P/IGF2R) Variants in American and Japanese Populations

J. Keith Killian,[1,2] Yoshihiko Oka,[1,3] Hong-Seok Jang,[1,4] Xialong Fu,[1,5] Robert A. Waterland,[1] Tetsuro Sohda,[6] Seigo Sakaguchi,[3] and Randy L. Jirtle[1,2]*

[1]*Department of Radiation Oncology, Duke University Medical Center, Durham, North Carolina*
[2]*Department of Pathology, Duke University Medical Center, Durham, North Carolina*
[3]*Department of Gastroenterology, Fukuoka University Chikushi Hospital, Fukuoka, Japan*
[4]*Department of Radiation Oncology, Uijongbo St. Mary's Hospital, Uijongbo, Korea*
[5]*Cancer Hospital/Cancer Institute, Shanghai Medical University, Shanghai, China*
[6]*Third Department of Internal Medicine, Fukuoka University School of Medicine, Fukuoka, Japan*

Communicated by Mark H. Paalman

M6P/IGF2R encodes a multifunctional protein involved in lysosomal enzyme trafficking, fetal organogenesis, tumor suppression, and cytotoxic T cell-induced apoptosis. M6P/IGF2R is imprinted and expressed only from the maternally inherited allele in marsupials and rodents. In contrast, humans were initially reported to differ from the imprinted mammalian orders by not having an imprinted M6P/IGF2R; however, some studies now suggest M6P/IGF2R imprinting may be a human polymorphic trait. Mutational and functional evidence are consistent with M6P/IGF2R also being a tumor suppressor in human colon, liver, lung, breast, and ovarian cancers. M6P/IGF2R expression is also pathologically downregulated following mammalian in vitro embryo culture, resulting in fetal overgrowth and "large offspring syndrome." Therefore, the M6P/IGF2R imprint status in humans is an unresolved question that critically impacts upon biological issues ranging from human cancer predisposition to evolution. Attempts to further characterize the imprint status of human M6P/IGF2R and loss of heterozygosity at this locus in cancer have been hindered by a lack of readily usable polymorphisms. To facilitate these genetic analyses, we have screened American and Japanese populations for M6P/IGF2R single nucleotide polymorphisms (SNPs). We have identified nine novel SNPs intragenic to human M6P/IGF2R, and have described experimental conditions for their optimal use. Three identified amino-acid variants in the M6P/IGF2R ligand-binding domains may be under selection in humans. Hum Mutat 18:25–31, 2001. © 2001 Wiley-Liss, Inc.

KEY WORDS: cancer; imprinting, genomic; fetal growth; human genetics; LOH; M6P; IGF2R; SNP; American; Japanese; population frequency

DATABASES:
IGF2R – OMIM: 147280; GDB: 120083; GenBank: Y00285 (cds), AF348209 (genomic);
http://www.GeneImprint.com (The Genomic Imprinting Website)

INTRODUCTION

The human *M6P/IGF2R* gene (*IGF2R*; MIM# 147280), located at 6q26-27 [Laureys et al., 1988; Rao et al., 1994], encodes for a multifunctional receptor that possesses distinct binding sites for several classes of molecules, including phosphomannosyl glycoproteins and insulin-like growth factor II (IGF2) [Jirtle,

Received 8 December 2000; accepted revised manuscript 30 March 2001.

*Correspondence to: Randy L. Jirtle, Departments of Radiation Oncology and Pathology, Duke University Medical Center, Durham, NC 27710. E-mail: jirtle@radonc.duke.edu

Contract grant sponsors: Sumitomo Chemical; AstraZeneca Pharmaceuticals; Contract grant sponsor: National Institutes of Health; Contract grant numbers: CA25951; ES08823; Contract grant sponsor: Department of Defense; Contract grant number: DAMD17-98-1-8305.

1999a]. The M6P binding site of this receptor mediates lysosomal enzyme trafficking, latent TGF-beta activation, and cell-mediated cytotoxic cell death [Jirtle, 1999b; Motyka et al., 2000]. M6P binding is the most ancient function attributed to this receptor, and it is present in invertebrates [Lakshmi et al., 1999], fish [Nadimpalli et al., 1999], amphibians and birds [Clairmont and Czech, 1989], and mammals [Killian et al., 2000]. The viviparous mammalian M6P/IGF2R also contains an independent binding site for IGF2; however, the marsupial homologue binds IGF2 with a 70-fold lower affinity than that of eutherian mammals. This difference in binding efficiency has been traced to repeat domain 11 of the receptor [Jirtle, 1999b; Killian et al., 2000]. Rather than provide an intracellular signal promoting cell growth, M6P/IGF2R internalizes and delivers bound IGF2 to the lysosomal compartment for degradation. Receptor acquisition of an IGF2 binding site occurred following the divergence of therian mammals (i.e. marsupials and true placental mammals) from the egg-laying monotreme mammals and non-mammalian vertebrates [Killian et al., 2000]. Thus, the history of M6P/IGF2R evolution reveals that its M6P binding function predates that for IGF2 by at least 300 million years.

M6P/IGF2R plays a critical role in embryonic development, immunity, and tumor suppression. M6P/IGF2R deficiency during murine development is associated with cardiac abnormalities, cleft palate, fetal overgrowth, and perinatal lethality [Lau et al., 1994; Ludwig et al., 1996; Wang et al., 1997; Melnick et al., 1998]. Recently, in an experimental model of "large offspring syndrome," Young et al. [2001] identified epigenetic changes intragenic to *M6P/IGF2R* that correlate with decreased gene expression and fetal overgrowth in livestock progeny which had been cultured in vitro during pre-implantation development. Consistent with the receptor's role in regulating cell growth and tissue proliferation, genetic analyses reveal frequent *M6P/IGF2R* loss of heterozygosity in liver, lung, breast, and ovarian cancer [De Souza et al., 1995; Hankins et al., 1996; Kong et al., 2000; Rey et al., 2000], and *M6P/IGF2R* is a common target of microsatellite instability in gastrointestinal tumors [Souza et al., 1996; Ouyang et al., 1997].

Interestingly, the M6P/IGF2R has been identified as a novel target of autoantibodies in patients with autoimmune diseases [Tarrago et al., 1999]. Furthermore, the M6P/IGF2R facilitates T cell activation by internalizing CD26/DPPIV (dipeptidyl peptidase IV), a cell surface T cell activation antigen [Ikushima et al., 2000]. Other recent findings indicate that Granzyme B internalization by M6P/IGF2R is required for cytotoxic T cells to induce apoptosis [Motyka et al., 2000]. The characterization of *M6P/IGF2R* allelic variants, including those that are non-synonymous and alter the protein product, will facilitate analysis of genetic linkage between immunopathologies and *M6P/IGF2R*.

M6P/IGF2R is expressed only from the maternally inherited allele in members of the marsupial [Killian et al., 2000] and rodent lineages [Barlow et al., 1991; Hu et al., 1998; Mills et al., 1998] while egg-laying animals are not imprinted at this locus [Killian et al., 2000; Nolan et al., 2001]. These observations are consistent with an ancestral mammalian origin of *M6P/IGF2R* imprinting roughly 150 million years ago; alternatively, *M6P/IGF2R* imprinting may have evolved convergently. Intriguingly, the human *M6P/IGF2R* is reported to have diverged from the imprinted mammalian orders such that individuals inherit two functional alleles [Kalscheuer et al., 1993; Ogawa et al., 1993; Smrzka et al., 1995; Riesewijk et al., 1996]; however, other reports suggest that imprinting at this locus is a polymorphic trait in humans with a minority population expressing only one copy of this gene [Xu et al., 1993; Xu et al., 1997; Riesewijk et al., 1998]. Polymorphic imprinting of *M6P/IGF2R* is predicted to predispose people to cancer because of the inheritance of a haploinsufficient tumor suppressor gene [Xu et al., 1997]. Recessive mutations and/or iatrogenic epimutations affecting embryonic and post-natal development would also show increased penetrance in imprinted individuals who have no recourse to a functional wild type allele.

The essential question of whether some humans inherit an imprinted *M6P/IGF2R* is still unresolved because allelic expression analysis is hindered by a dearth of characterized polymorphisms within the coding sequence. We have herein identified nine novel human *M6P/IGF2R* polymorphisms to further facilitate such studies, and have compiled them into a table with those previously characterized. We also have described experimental conditions for their op-

timal use, and have discussed those polymorphic variants prone to artifactual results.

MATERIALS AND METHODS

DNA Samples

Blood samples for DNA genotyping were obtained from healthy adults at Duke University Medical Center in the United States and Fukuoka University Hospital in Japan. Use of these samples was approved by the Duke University Medical Center Institutional Review Board. Genomic DNA was extracted from blood according to the manufacturer's protocol (QIAamp DNA Blood Kit, QIAGEN Inc., Valencia, CA).

Human M6P/IGF2R SNP Identification

The 48 individual *M6P/IGF2R* exons with approximately 100 bp flanking intronic sequence were PCR-amplified from 12 individuals using oligonucleotides previously described [Killian and Jirtle, 1999]. PCRs were performed with 50–100 ng template DNA, 1.5 U Platinum Taq DNA polymerase (Life Technologies, Baltimore, MD), 15 pmol primers, 1.5 mM $MgCl_2$ and 100 μM dNTPs in a 30 μL PCR reaction volume (94°C × 15 sec, 55°C × 5 sec, and 72°C × 45 sec for 30–35 cycles). Amplimers were analyzed and extracted from 2% agarose gel (GenElute, Sigma, St. Louis, MO), and sequenced either manually (Thermo Sequinase, USB, Cleveland, OH) or on the ABI Prism 377 using BigDye terminators (PE Biosystems, Foster City, CA). SNPs were identified as sequence dimorphisms in the sequenced DNA or di-allelic homozygous alternates.

Human M6P/IGF2R SNP Analysis

PCR primers for amplification of the 10 *M6P/IGF2R* SNPs are given in Table 1. The allele frequencies of these 10 polymorphisms were determined in a total of 93 individuals, including 43 American (25 white American, 16 black American, and two Asian American) and 50 Japanese individuals. PCRs were performed with 50–100 ng template DNA, 1.5 U Platinum Taq DNA polymerase, 15 pmol primers, 1.5 mM $MgCl_2$ and 100 μM dNTPs in a 30 μL PCR reaction volume (94°C × 15 sec, 55°C × 5 sec, and 72°C × 45 sec for 30–35 cycles). Genotypes were confirmed by DNA sequencing. The distribution of allelic variants among American and Japanese populations was analyzed by chi-square tests. To decrease the likelihood of alpha-error associated with multiple testing, only differences with $P<0.01$ were considered statistically significant.

RESULTS

The population-specific frequency of one previously identified [Zhong et al., 1999], and nine novel *M6P/IGF2R* single nucleotide polymorphisms (SNPs) are provided in Table 2. Allele frequencies were derived from analysis of 50 Japanese and 43 Americans (16 black, 25 white, two Asian). No statistically significant differences in allelic distribution were found between black and white Americans, so these populations were pooled in subsequent analyses. Comparing the American and Japanese samples, highly significant differences in allelic distribution were found for six of the 10 SNPs. The polymorphisms c.1197A>G, c.1737G>A, and c.2286G>A, have been described as silent mutations in ovarian cancer [Rey et al., 2000]. Our analysis shows that these variants represent common polymorphisms in the human population, and suggests no particular relationship to ovarian cancer. Two of the 10 polymorphisms are transversions (c.901C>G and IVSX47–5A>T), while the remainder are transitions.

DISCUSSION

We have identified novel allelic variants of the human *M6P/IGF2R*, and have shown that

TABLE 1. PCR Primers for Analysis of Human *M6P/IGF2R* SNPs

SNP	Location	F primer	R primer
c. 901C>G	Exon 6	CTAAGGGTACTGTGATTATCACTC	GAAAGTCAGGTCCTTGCTGGAG
IVS8-22G>T	Intron 7	GTGGAAAATCTGCATTAAGCTGCATG	CCTTCTTCCTAAGCAGCGCC
c. 1197A>G	Exon 9	GACTAAGTAAGACTGTAATCTTCTAATACC	CGCACAGAGGTTGTTGACGTAC
c. 1737A>G	Exon 12	GTGACTCAGAGAAATGAGCATTGC	CTAACTCATTCCAAACTGGATGCC
c. 2286A>G	Exon 16	GTGACTCCTCACGTCGCTCACG	CACAGGCATGAGTATCCTCAGG
IVS23-42C>T	Intron 22	CTGCACTGTGCTTGTGGGCTGC	GACTCTTGACCGGCCTCTCAGTTC
c. 5002A>G	Exon 34	GAAATTGATGGTCCTGACTTGCG	GCACTGGAGATGCACTTCTCC
c. 6206A>G	Exon 40	GCATAGACACAGTGACAGTCTGATC	GCAGTCTGAAGTTCACATGC
IVS47-107A>G	Intron 46	CCATGCCCTCTCTACACTGGAG	CCTGATGAGAACGACATGGACAGC
IVS47-5A>T	Intron 46	CCATGCCCTCTCTACACTGGAG	CCTGATGAGAACGACATGGACAGC

TABLE 2. Allelic Variants of Human M6P/IGF2R

Variant	Amino acid change	Allele	American			Japanese			P value
			n	(%)	Frequency	n	(%)	Frequency	
c. 901C>G	CTG>GTG	CC	32	(74.4)	C=0.8721	28	(56.0)	C=0.7500	0.086
(Exon 6)	Leu>Val	CG	11	(25.6)	G=0.1279	19	(38.0)	C=0.2500	
		GG	0	(0)		3	(6.0)		
IVS8-22G>T		GG	1	(2.3)	G=0.1977	12	(24.0)	G=0.5100	7×10^{-5}
(Intron 7)		GT	15	(34.9)	T=0.8023	27	(54.0)	T=0.4900	
		TT	27	(62.8)		11	(22.0)		
c. 1197A>G	TCA>TCG	AA	24	(55.8)	A=0.6744	2	(4.0)	A=0.2400	2×10^{-7}
(Exon 9)	Ser	AG	10	(23.3)	G=0.3256	20	(40.0)	G=0.7600	
		GG	9	(20.9)		28	(56.0)		
c. 1737A>G	GGA>GGG	AA	7	(16.3)	A=0.3023	30	(60.0)	A=0.7800	5×10^{-8}
(Exon 12)	Gly	AG	12	(27.9)	G=0.6977	18	(36.0)	G=0.2200	
		GG	24	(55.8)		2	(4.0)		
c. 2286A>G	ACA>ACG	AA	12	(27.9)	A=0.5116	3	(6.0)	A=0.2500	0.002
(Exon 16)	Thr	AG	20	(46.5)	G=0.4884	19	(38.0)	G=0.7500	
		GG	11	(25.6)		28	(56.0)		
IVS23-42C>T		CC	2	(5.0)	C=0.2250	7	(14.0)	C=0.3500	0.203
(Intron 22)		CT	14	(35.0)	T=0.7750	21	(42.0)	T=0.6500	
		TT	24	(60.0)		22	(44.0)		
c. 5002A>G	AGG>GGG	AA	0	(0)	A=0.1512	1	(2.0)	A=0.1700	0.647
(Exon 34)	Arg>Gly	AG	13	(30.2)	G=0.8488	15	(30.0)	G=0.8300	
		GG	30	(69.8)		34	(68.0)		
c. 6206A>G	AAC>AGC	AA	38	(90.5)	A=0.9524	22	(44.0)	A=0.6500	1×10^{-5}
(Exon 40)	Asn>Ser	AG	4	(9.5)	G=0.0476	21	(42.0)	G=0.3500	
		GG	0	(0)		7	(14.0)		
IVS47-107A>G		AA	20	(47.6)	A=0.6429	6	(12.0)	A=0.4000	8×10^{-4}
(Intron 46)		AG	14	(33.3)	G=0.3571	28	(56.0)	G=0.6000	
		GG	8	(19.1)		16	(32.0)		
IVS47-5A>T		AA	5	(11.6)	A=0.2558	8	(16.0)	A=0.4000	0.059
(Intron 46)		AT	12	(27.9)	T=0.7442	24	(48.0)	T=0.6000	
		TT	26	(60.5)		18	(36.0)		

Polymorphisms in the coding region and introns are numbered according to Morgan et al. [1987], with nomenclature based on Antonarakis [1998]. Intronic polymorphisms are amplified with the published primers for the nearest exon [Killian and Jirtle, 1999]. Note the presence of non-synonymous polymorphisms in exon 6 (c. 901C>G, Leu252Val); exon 34 (c. 5002G>A, Gly1619Arg); and exon 40 (c. 6206A>G, Asn2020Ser). *P* value refers to the significance of statistical comparison between American and Japanese allele frequencies at each SNP locus.

they are present in both the American (black and white) and Japanese populations, although at different frequencies (Table 2). Because these polymorphisms are shared by these populations, they were possibly present prior to the geographic separation of ancestors common to blacks, whites, and Japanese. Alternatively, the variants could have originated after the divergence of these populations with subsequent introgression. Because of the abundance of SNPs with population-specific frequencies, the *M6P/IGF2R* may serve as a model gene for future analyses of human relatedness and diasporas. Comparison of the combined American versus Japanese populations reveals that six out of 10 SNPs have significantly different allele frequencies ($P<0.01$). These SNPs include IVS8-22G>T, c.1197A>G, c.1737A>G, c.2286A>G, c.6206A>G, and IVS47-107A>G.

Interestingly, the *M6P/IGF2R* variant manifesting the least difference between Americans and Japanese is the non-synonymous c.5002A>G. The American and Japanese populations analyzed in this study both contained 30% heterozygosity at this locus. However, when we split the American population by race, we noted a discrepancy in c.5002A>G heterozygote frequency, with 12.5% and 36% heterozygosity in black and white, respectively ($P=0.098$). Analysis of native African and European populations would help elucidate whether c.5002A>G is ancestral to humans, or whether it has arisen convergently in one or more populations, with subsequent introgression into the black American gene pool.

It is striking that three out of six of the polymorphisms we identified within the human *M6P/IGF2R* coding region are non-synonymous. While c.901C>G, Leu252Val is a relatively conservative amino acid change, c.5002G>A, Gly1619Arg, and c.6206A>G, Asn2020Ser variants differ substantially. The c.901C>G, Leu252Val resides in repeat domain 2 of the

M6P/IGF2R protein, a domain reported to be involved in M6P binding [Jirtle, 1999b]. The c.5002G>A, Gly1619Arg falls within repeat domain 11 of the protein, which contains the IGF2 binding site [Jirtle, 1999b]. Finally, c.6206A>G, Asn2020Ser is within repeat domain 13, which reportedly is an IGF2 affinity-enhancing domain [Jirtle, 1999b].

Although c.901C>G, Leu252Val (exon 9) is not a radical alteration, there is strict conservation of leucine at this position in the alignment of platypus, opossum, mouse, rat, and cow M6P/IGF2R [Killian et al., 2000]. Because this amino acid variation occurs in the M6P binding region, it may affect functions including lysosomal enzyme trafficking, TGF-β activation, and cytotoxic T-cell mediated tumor cell apoptosis.

Concerning c.5002G>A, Gly1619Arg, we also note strict conservation of glycine at this position in platypus, opossum, mouse, rat, and cow [Killian et al., 2000], suggesting that Arg1619 may indeed have a functional consequence in humans. Because of its positioning in the IGF2-binding repeat domain 11, the functional consequence of inheriting a heterozygous Gly1619Arg M6P/IGF2R complement may affect IGF2 binding. Because fetal growth is largely controlled by M6P/IGF2R-mediated IGF2 degradation, the IGF2-binding domain variants of human M6P/IGF2R are potentially involved in offspring size in humans. The possible involvement of c.5002G>A in human pathology requires further linkage and functional studies. Interestingly, Glycine 1619 is conserved in platypus, an animal whose M6P/IGF2R reportedly does not bind IGF2 [Killian et al., 2000].

The third non-synonymous polymorphism in the *M6P/IGF2R* coding sequence, c.6206A>G, Asn2020Ser (exon 40), occurs at a relatively variable position in the mammalian comparison [Killian et al., 2000]: platypus and opossum have an aspartic acid, mouse and rat a glutamic acid, and cow an asparagine. Elucidating the roles of these protein-altering M6P/IGF2R variants in human traits requires biochemical investigation.

Experimentally, while most *M6P/IGF2R* polymorphisms clearly distinguished the parental alleles, there are some precautions. The c.2286A>G and c.5002G>A polymorphic sites in exons 16 and 34, respectively, show band compressions when sequenced manually with radio-labeled terminators, i.e., Thermo Sequenase. This requires the use of "GC-rich" sequencing protocols, including substituting deoxyinosine for deoxyguanosine when DNA sequencing. These compression artifacts are not observed using the ABI automated sequencer with BigDye terminators.

There are two previously described polymorphisms in the distal end of the 3′ UTR that are particularly problematic for PCR analysis: a GT dinucleotide repeat polymorphism [Goto et al., 1992] and a tetranucleotide insertion/deletion [Hol et al., 1992]. The GT repeat is prone to the usual Taq DNA polymerase slippage artifacts during PCR, and the results can be further obfuscated by the presence or absence of the tetranucleotide insertion/deletion that is only 27 bp away. The PCR primers initially reported by Goto et al. [1992], and utilized in some studies of human *M6P/IGF2R* imprinting, in fact hybridize directly to the tetranucleotide insertion/deletion, such that primer binding is dependent upon the status of the ACAA insertion/deletion polymorphism. Results are therefore often difficult to interpret when the 3′ UTR GT-repeat polymorphism is used for parental allelic analysis. In contrast, we invariably obtain unambiguous and reproducible results when samples are characterized using the SNPs reported herein.

Regardless of which polymorphism is employed, we have observed frequent stochastic amplification of one variant over another whenever more than one round of PCR and/or greater than 35 cycles in a single round of PCR are used. If insufficient starting template is available to yield a robust amplimer in a single round of PCR, results must be confirmed by several repeat analyses, and preferably with more than one polymorphism.

To summarize our experience with *M6P/IGF2R* allelic analysis, we recommend using SNPs rather than the GT-repeat polymorphism for *M6P/IGF2R* allelic expression and LOH analyses. Starting template of sufficient quality and concentration should be used to ensure robust amplification in less than 35 total PCR cycles. In the case of small quantities of tissue or paraffin-embedded tissues in which this is not always possible, repeat analyses must be performed to rule out stochastic amplification. It is important to note that of the published papers addressing the issue of *M6P/IGF2R* imprinting in humans, deviation from biallelic expression has only been reported when using the ACAA

insertion/deletion and GT-repeat polymorphisms [Kalscheuer et al., 1993; Ogawa et al., 1993; Riesewijk et al., 1996; Smrzka et al., 1995]. Whether these reports will be supported by expression studies using these novel SNPs remains to be determined.

In conclusion, our identification and characterization of nine novel *M6P/IGF2R* SNPs is essential for researchers to address several questions pertaining to human development, cancer, and imprinting. These include: 1) determining whether some humans inherit a cancer-predisposing, imprinted *M6P/IGF2R*; 2) evaluating the role of *M6P/IGF2R* allelic loss in cancer patient prognosis and response to therapy; 3) determining the *M6P/IGF2R* mutation status in cancers from different ethnic populations; and 4) assessing the potential influence of human embryo culture procedures such as IVF on *M6P/IGF2R* expression. The availability of abundant, high-quality *M6P/IGF2R* polymorphisms will enable these important questions to be rapidly and conclusively resolved. Finally, just as *M6P/IGF2R* is a model gene for mammalian phylogenetic inference [Killian et al., 2001], we expect population analyses of *M6P/IGF2R* variants to be highly informative for studies of human evolution, relatedness, and diasporas.

ACKNOWLEDGMENTS

We thank the Duke DNA Analysis Facility for genomic sequencing, and Mark Paalman for helpful comments on improving the manuscript.

REFERENCES

Antonarakis SE, Nomenclature Working Group. 1998. Recommendations for a nomenclature system for human gene mutations. Hum Mutat 11:1–3.

Barlow DP, Stöger R, Herrmann BG, Saito K, Schweifer N. 1991. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. Nature 349:84–87.

Clairmont KB, Czech MP. 1989. Chicken and Xenopus mannose 6-phosphate receptors fail to bind insulin-like growth factor II. J Biol Chem 264:16390–16392.

De Souza AT, Hankins GR, Washington MK, Orton TC, Jirtle RL. 1995. M6P/IGF2R gene is mutated in human hepatocellular carcinomas with loss of heterozygosity. Nat Genet 11:447–449.

Goto J, Figlewicz DA, Marineau C, Khodr N, Rouleau GA. 1992. Dinucleotide repeat polymorphism at the IGF2R locus. Nucleic Acids Res 20:923.

Hankins GR, De Souza AT, Bentley RC, Patel MR, Marks JR, Iglehart JD, Jirtle RL. 1996. *M6P/IGF2* receptor: a candidate breast tumor suppressor gene. Oncogene 12:2003–2009.

Hol FA, Geurds MP, Hamel BC, Mariman EC. 1992. Improving the polymorphism content of the 3′ UTR of the human IGF2R gene. Hum Mol Genet 1:347.

Hu JF, Oruganti H, Vu TH, Hoffman AR. 1998. Tissue-specific imprinting of the mouse insulin-like growth factor II receptor gene correlates with differential allele-specific DNA methylation. Mol Endocrinol 12:220–232.

Ikushima H, Munakata Y, Ishii T, Iwata S, Terashima M, Tanaka H, Schlossman SF, Morimoto C. 2000. Internalization of CD26 by mannose 6-phosphate/insulin-like growth factor II receptor contributes to T cell activation. Proc Natl Acad Sci USA 97:8439–8444.

Jirtle RL. 1999a. Genomic imprinting and cancer. Exp Cell Res 248:18–24.

Jirtle RL. 1999b. Mannose 6-phosphate receptors. In: Creidton TE, editor. Encyclopedia of molecular biology. New York: Wiley-Liss, Inc. p 1441–1447.

Kalscheuer VM, Mariman EC, Schepens MT, Rehder H, Ropers H-H. 1993. The insulin-like growth factor type-2 receptor gene is imprinted in the mouse but not in humans. Nat Genet 5:74–78.

Killian JK, Jirtle RL. 1999. Genomic structure of the human *M6P/IGF2* receptor. Mamm Genome 10:74–77.

Killian JK, Byrd JC, Jirtle JV, Munday BL, Stoskopf MK, MacDonald RG, Jirtle RL. 2000. M6P/IGF2R imprinting evolution in mammals. Mol Cell 5:707–716.

Killian JK, Buckley T, Stewart N, Munday B, Jirtle RL. 2001. Marsupials and eutherians reunited. Genetic evidence for the theria hypothesis of mammalian evolution. Mamm Genome, in press.

Kong FM, Anscher MS, Washington MK, Killian JK, Jirtle RL. 2000. M6P/IGF2R is mutated in squamous cell carcinoma of the lung. Oncogene 19:1572–1578.

Lakshmi YU, Radha Y, Hille-Rehfeld A, von Figura K, Kumar NS. 1999. Identification of the putative mannose 6-phosphate receptor protein (MPR 300) in the invertebrate unio. Biosci Rep 19:403–409.

Lau MM, Stewart CE, Liu Z, Bhatt H, Rotwein P, Stewart CL. 1994. Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. Genes Dev 8:2953–2963.

Laureys G, Barton DE, Ullrich A, Francke U. 1988. Chromosomal mapping of the gene for the type II insulin-like growth factor receptor/cation-independent mannose 6-phosphate receptor in man and mouse. Genomics 3:224–229.

Ludwig T, Eggenschwiler J, Fisher P, D'Ercole AJ, Davenport ML, Efstratiadis A. 1996. Mouse mutants lacking the type 2 IGF receptor (IGF2R) are rescued from perinatal lethality in Igf2 and Igf1r null backgrounds. Dev Biol 177:517–535.

Melnick M, Chen H, Buckley S, Warburton D, Jaskoll T. 1998. Insulin-like growth factor II receptor, transforming growth factor-beta, and Cdk4 expression and the developmental epigenetics of mouse palate morphogenesis and dysmorphogenesis. Dev Dyn 211:11–25.

Mills JJ, Falls JG, De Souza AT, Jirtle RL. 1998. Imprinted *M6p/Igf2* receptor is mutated in rat liver tumors. Oncogene 16:2797–2802.

Morgan DO, Edman JC, Standring DN, Fried VA, Smith MC, Roth RA, Rutter WJ. 1987. Insulin-like growth factor II receptor as a multifunctional binding protein. Nature 329:301–307.

Motyka B, Korbutt G, Pinkoski MJ, Heilbein JA, Caputo A, Hobman M, Barry M, Shostak I, Sawchuk T, Holmes CF, Gauldie J, Bleackley RC. 2000. Mannose 6-phosphate/insulin-like growth factor II receptor is a death receptor for granzyme B during cytotoxic T cell-induced apoptosis. Cell 103:491–5000.

Nadimpalli SK, Yerramalla UL, Hille-Rehfeld A, von Figura K. 1999. Mannose 6-phosphate receptors (MPR 300 and MPR 46) from a teleostean fish (trout). Comp Biochem Physiol B Biochem Mol Biol 123:261–265.

Nolan CM, Killian JK, Petitte JN, Jirtle RL. 2001. Imprint status of M6P/IGF2R and IGF2 in chickens. Genes Evol Dev, in press.

Ogawa O, McNoe LA, Eccles MR, Morison IM, Reeve AE. 1993. Human insulin-like growth factor type I and type II receptors are not imprinted. Hum Mol Genet 2:2163–2165.

Ouyang H, Shiwaku HO, Hagiwara H, Miura K, Abe T, Kato Y, Ohtani H, Shiiba K, Souza RF, Meltzer SJ, Horii A. 1997. The insulin-like growth factor II receptor gene is mutated in genetically unstable cancers of the endometrium, stomach, and colorectum. Cancer Res 57:1851–1854.

Pulford DJ, Falls JG, Killian JK, Jirtle RL. 1999. Polymorphisms, genomic imprinting and cancer susceptibility. Mutat Res 436:59–67.

Rao PH, Murty VV VS, Gaidano G, Hauptschein R, Dalla-Favera R, Chaganti RSK. 1994. Subregional mapping of 8 single copy loci to chromosome 6 by fluorescence in situ hybridization. Cytogenet Cell Genet 66:272–273.

Rey JM, Theillet C, Brouillet JP, Rochefort H. 2000. Stable amino-acid sequence of the mannose-6-phosphate/insulin-like growth-factor-II receptor in ovarian carcinomas with loss of heterozygosity and in breast-cancer cell lines. Int J Cancer 85:466–473.

Riesewijk AM, Schepens MT, Welch TR, Van den Berg-Loonen EM, Mariman EM, Ropers HH, Kalscheuer VM. 1996. Maternal-specific methylation of the human *IGF2R* gene is not accompanied by allele-specific transcription. Genomics 31:158–166.

Riesewijk AM, Xu YQ, Schepens MT, Mariman EM, Polychronakos C, Ropers HH, Kalscheuer VM. 1998. Absence of an obvious molecular imprinting mechanism in a human fetus with monoallelic IGF2R expression. Biochem Biophys Res Commun 245:272–277.

Smrzka OW, Fae I, Stoger R, Kurzbauer R, Fischer GF, Henn T, Weith A, Barlow DP. 1995. Conservation of a maternal-specific methylation signal at the human IGF2R locus. Hum Mol Genet 4:1945–1952.

Souza RF, Appel R, Yin J, Wang S, Smolinski KN, Abraham JM, Zou T-T, Shi Y-Q, Lei J, Cottrell J, Cymes K, Biden K, Simms L, Leggett B, Lynch PM, Frazier M, Powell SM, Harpaz N, Sugimura H, Young J, Meltzer SJ. 1996. The insulin-like growth factor II receptor gene is a target of microsatellite instability in human gastrointestinal tumours. Nat Genet 14:255–257.

Tarrago D, Aguilera I, Melero J, Wichmann I, Nunez-Roldan A, Sanchez B. 1999. Identification of cation-independent mannose 6-phosphate receptor/insulin-like growth factor type-2 receptor as a novel target of autoantibodies. Immunology 98:652–662.

Wang SN, Souza RF, Kong DH, Yin J, Smolinski KN, Zou TT, Frank T, Young J, Flanders KC, Sugimura H, Abraham JM, Meltzer SJ. 1997. Deficient transforming growth factor-β1 activation and excessive insulin-like growth factor II (IGFII) expression in IGFII receptor-mutant tumors. Cancer Res 57:2543–2546.

Xu Y, Goodyer CG, Deal C, Polychronakos C. 1993. Functional polymorphism in the parental imprinting of the human IGF2R gene. Biochem Biophys Res Commun 197: 747–754.

Xu YQ, Grundy P, Polychronakos C. 1997. Aberrant imprinting of the insulin-like growth factor II receptor gene in Wilms' tumor. Oncogene 14:1041–1046.

Young LE, Fernandes K, McEvoy TG, Butterwith SC, Gutierrez CG, Carolan C, Broadbent PJ, Robinson JJ, Wilmut I, Sinclair KD. 2001. Epigenetic change in IGF2R is associated with fetal overgrowth after sheep embryo culture. Nat Genet 27:153–154.

Zhong X, Hemmi H, Shimatake H. 1999. A common polymorphism in exon 40 of the human mannose 6- phosphate/insulin-like growth factor II receptor gene. Mol Cell Probes 13:397–400.